AIMS Mathematics

https://www.aimspress.com/journal/Math

AIMS Mathematics, 10(12): 30271–30306.
DOI: 10.3934/math.20251330
Received: 15 October 2025
Revised: 05 December 2025
Accepted: 17 December 2025
Published: 24 December 2025

Research article

Fixed-time optimal consensus for nonlinear strict-feedback multi-agent systems based on reinforcement learning and neural network observers

Kaile Zhang[1], **Zhanheng Chen**[1,2,*], **Zhiyong Yu**[3] **and Haijun Jiang**[3]

[1] College of Mathematics and Statistics, Yili Normal University, Yining 835000, China
[2] Institute of Applied Mathematics, Yili Normal University, Yining 835000, China
[3] College of Mathematics and System Sciences, Xinjiang University, Urumqi 830017, China

* **Correspondence:** Email: zhchen@ylnu.edu.cn.

Abstract: This paper investigates the fixed-time consensus control problem for strict-feedback multi-agent systems based on reinforcement learning. First, under the observer–critic–actor framework, neural networks are applied to the observer to address the issue of unmeasurable system states and nonlinear functions. Furthermore, based on the backstepping method, a reinforcement learning algorithm is constructed to obtain the optimal control input, which is then evaluated and optimized by the critic–actor network to derive an approximate optimal control input. Second, by constructing a Lyapunov function and utilizing the boundedness of the critic–actor network matrix trace along with Lyapunov stability theory, the fixed-time consensus of the system is proven. Finally, the effectiveness of the algorithm is verified through numerical simulations.

Keywords: neural network; reinforcement learning; fixed-time consensus; backstepping method; critic–actor network
Mathematics Subject Classification: 93A14, 93D50, 90C39, 93E11

1. Introduction

For decades, the study of multi-agent systems (MASs) has been a major focus of academic research. As a subfield of MASs, multi-agent cooperative control has found extensive applications in spacecraft formation flight [1], smart grids [2], and vehicle systems [3,4]. Consensus control is the core of multi-agent cooperative control. Its main purpose is to achieve the unification and coordination of each agent within the group under specific states or attributes, thereby enabling better communication among the agents. With the continuous deepening of research, certain specific types of MASs, such as nonlinear strict-feedback MASs, have gained significant attention due to their strong robustness and distributed computing capabilities. However, the nonlinear functions in these systems are often unknown. On the other hand, neural networks (NNs) and fuzzy logic systems (FLSs) demonstrate unique advantages in approximating unknown functions due to their powerful nonlinear mapping capabilities, self-learning characteristics, and strong generalization abilities. Thus, in recent years, the use of NNs or FLSs to approximate unknown nonlinear functions in strict-feedback systems has attracted growing research interest. In the literature [5], NNs are used for the construction of observers and the approximation of nonlinear functions. The literature [6] established sufficient conditions for achieving consensus in MASs without Zeno behavior by combining neural network technology with backstepping control. In literature [7, 8], for nonlinear multi-agent systems subject to output constraints and denial-of-service (DoS) attacks, the problem of unknown nonlinear functions is addressed using fuzzy logic technology, and it is proven that the system can achieve an adaptive fault-tolerant consensus. Furthermore, if a control algorithm can be designed to reduce the convergence time of a strict-feedback multi-agent system to a desired value, the system will then be able to stabilize near the equilibrium point within a preset time. Subsequently, this complex nonlinear strict-feedback system will demonstrate higher convergence accuracy and stronger anti-disturbance capability.

Although finite-time (FT) consensus can ensure that the system converges to a stable state within a predetermined time, its convergence time depends on the initial values of the system [9, 10]. On the one hand, the convergence time may significantly increase as the initial state deviates further from the equilibrium point [11], This may prevent the system from meeting the predefined rapid-response requirements, especially when the initial deviation is large. Moreover, if the initial state of the system is unavailable, it becomes impossible to accurately estimate the convergence time or ensure that the system stabilizes within the preset duration [12]. To address these challenges, Polyakov introduced the concept of fixed-time (FxT) stability in [13]. Compared with finite-time stability, its advantage lies in that the convergence time of the system is determined solely by the control input and is independent of the initial state. The literature [14] investigated adaptive FxT optimal formation control for uncertain nonlinear multi-agent systems, while [15] studied the FxT consensus problem for a class of second-order nonlinear multi-agent systems under actuator faults, proposing a fully distributed controller that enables the considered MASs to achieve consensus. In [16], a FxT state observer was employed to estimate the leader's state, and a nonsingular terminal sliding mode adaptive controller was designed to enhance system robustness and improve tracking accuracy. In [17], an edge-based adaptive control protocol was proposed to solve the FxT consensus control problem for nonlinear leader–follower multi-agent systems. This protocol allows for more accurate assignment of time-varying weights to different communication edges. Research on FxT consensus control is often combined with systems such as leader-follower multi-agent systems, second-order nonlinear multi-agent systems, and heterogeneous multi-agent systems [18]. However, achieving a consensus achievement of strict–feedback multi-agent systems relies on accurate feedback of the neighbors' states, while FxT control requires precise global synchronization as a prerequisite. Therefore, the integration of FxT consensus control with strict-feedback multi-agent systems remains an underexplored research area. In light of this gap, establishing a FxT consensus for such systems carries substantial theoretical importance and practical relevance. If a more powerful control input can be obtained, it can not only shorten the system's convergence time but also reduce the deviations in the convergence states of each agent. In existing research, the acquisition of such optimal control inputs is usually realized by constructing reinforcement learning (RL) algorithms.

Optimal control based on RL was first proposed by Werbos in the literature [19]. The application of

RL algorithms to strict–feedback systems has since become a hotspot in control theory research [20–23]. In [23], for nonlinear strict–feedback multi-agent systems with actuator faults, an adaptive RL algorithm was proposed to effectively compensate for the impact of such faults. In [24], for a class of nonlinear strict–feedback systems with unknown functions, the backstepping method was integrated into the RL algorithm. By constructing virtual controllers and actual controllers, the global optimal control of the system was achieved. In [25], an adaptive RL algorithm was developed for strict–feedback multi-agent systems with time-varying bias. This algorithm was specifically designed to compensate for the lack of state information and obtain control inputs, thereby achieving the optimal control. In [26], an observer and an error derivative-based objective function were constructed, leading to an improved optimal backstepping control method. This approach resolves the issue where the optimal control input derived from a cost function with exponential factors fails to ensure asymptotic error stability. In [27], a RL algorithm was proposed to address the H_∞ consensus problem for discrete-time multi-agent systems under DoS attacks. The algorithm can obtain the target gain matrix without requiring knowledge of the system's dynamics. Numerical simulations demonstrate that in the presence of actuator faults or DoS attacks, the designed RL algorithm can mitigate their impacts on the multi-agent system and ensure the smooth achievement of consensus. RL algorithms have also been applied to reduce systems' convergence time. For example, applying RL algorithms to finite-time consensus control in MASs helps shorten the time required for the systems to achieve convergence [28, 29]. However, as previously discussed, FT control is limited by its dependence on the system's initial conditions. Therefore, combining RL algorithms with FxT consensus control for MASs is of significant research importance. On the basis of ensuring that "the system is independent of initial states", it can further reduce quality deviations, lower control energy consumption, and thereby achieve the dual goals of both fast convergence and optimal control.

In summary, this paper investigates the FxT consensus problem for strict-feedback multi-agent systems by integrating a reinforcement learning algorithm with neural network observers. By using observer–critic–actor networks to obtain the optimal control input, the time required for the system to achieve FxT convergence is shortened. The main contributions are as follows.

(1) Unlike the critic–actor network frameworks proposed in [23–25, 27] that primarily focus on evaluating and tuning the derived optimal control input, this paper analyzes the conditions for the existence of a minimum residual in the Hamilton–Jacobi–Bellman (HJB) equation. This analysis provides a basis for enhanced evaluation and adjustment of the optimal control input.

(2) Within this critic–actor framework, by introducing an unknown nonlinear function to determine the dimension of the framework, the originally high-complexity vector operations can be effectively reduced to scalar computations. This transformation further enables the derivation of an upper bound for the trace of the critic–actor network weight matrix. Compared with the methods in [18, 23–25, 29] that do not account for the network's dimensions, the weight matrix obtained by our method achieves higher stability.

(3) For this RL algorithm, we have developed an objective function that incorporates both the error and its derivative terms. By constructing a Lyapunov function and performing differentiation. This function demonstrates greater stability than those in [18, 24, 29]. Subsequently, numerical simulations verify that the optimal control input derived from the algorithm exhibits high stability. Furthermore, the process of proving FxT stability validates that the proposed algorithm can shorten the convergence time for MASs to achieve FxT stability.

Notations. The following notations are used throughout this paper. Let R denote the set of real numbers, $\mathbb{R}^n$ stands for the n-dimensional Euclidean space, and $\mathbb{R}^{N\times N}$ represents the $N\times N$ dimensional real matrices. Let A^T represents the transpose of the matrix A. I_n is an n-order identity matrix, $0_N \in \mathbb{R}^{N\times N}$ is a vector with all the entries being 0, diag$\{\ldots\}$ represents a diagonal matrix, and 1_N represents $(1, 1, \ldots, 1)^T$. $\lambda_{\min}(M)$ represents the minimum eigenvalues of a symmetric matrix M.

2. Preliminaries and model description

2.1. Graph theory knowledge

The communication topology of a MASs is described by an undirected graph $G = (V, E, A)$, where $V = \{v_1, v_2, \ldots, v_n\}$ is the set of nodes, $E \subseteq V \times V$ is the set of edges between any two nodes, and $A = [a_{ij}]_{n\times n}$ is the adjacency matrix of the system. If $(v_i, v_j) \in E$, then there is an edge connecting node v_i and node v_j, denoted as ε_{ij}. When $\varepsilon_{ij} \in E$ with $i \neq j$, we have $a_{ij} = 1$; otherwise, $a_{ij} = 0$. Since G is undirected, the adjacency matrix is symmetric; that is, $a_{ij} = a_{ji}$ for all i, j. The Laplacian matrix of the system is defined as $L = \Delta - A$, where Δ is the degree matrix, given by $\Delta = \text{diag}\left(\sum_{j=1}^{n} a_{1j},\ \sum_{j=1}^{n} a_{2j},\ \ldots,\ \sum_{j=1}^{n} a_{nj}\right)$. Since the graph under consideration is undirected and connected, the associated Laplacian matrix L is symmetric and positive semi-definite. A diagonal matrix $D = \text{diag}(d_1, d_2, \ldots, d_n)$ is used to characterize the information exchange between the followers and the leader. Specifically, $d_i = 1$ if the i-th follower exchanges information with the leader; otherwise, $d_i = 0$.

2.2. System model

Consider the following nonlinear strict-feedback system:

$$\begin{cases} \dot{x}_{i,j}(t) = x_{i,j+1}(t) + f_{i,j}(\bar{x}_{i,j}(t)), \\ \dot{x}_{i,n}(t) = u_i(t) + f_{i,n}(\bar{x}_{i,n}(t)), \\ y_i(t) = x_{i,1}(t), \\ i = 1, 2, \cdots, N, j = 1, 2, \cdots, n-1, \end{cases} \tag{2.1}$$

where $\bar{x}_{i,j}(t) = [x_{i,1}(t), \cdots, x_{i,j}(t)]^T \in \mathbb{R}^j$ and $\bar{x}_{i,n}(t) = [x_{i,1}(t), \cdots, x_{i,n}(t)]^T \in \mathbb{R}^n$ denote the state vectors of the system. Except for $x_{i,1}(t)$, all other system states are unmeasurable. Here, $f_{i,j}(\bar{x}_{i,j})$ $(i = 1, 2, \cdots, N;\ j = 1, 2, \cdots, n)$ represents the nonlinear functions of the system, $u_i(t)$ denotes the control input of the system, and $y_i(t)$ stands for the control output of the i-th follower.

Assumption 2.1. *[25] The output of the leader in the system, $y_d(t)$, and its arbitrary-order derivative $y_d^{(r)}(t)$ are both continuous and bounded.*

Lemma 2.1. *[30] For arbitrary $w_1, w_2, \cdots, w_n \in \mathbb{R}$, $0 < p < 1$, and $1 < q < +\infty$, the following inequality holds:*

$$\begin{cases} (\sum_{i=1}^{n} |w_i|)^p \leq \sum_{i=1}^{n} |w_i|^p \leq n^{1-p} (\sum_{i=1}^{n} |w_i|)^p, \\ n^{1-q} (\sum_{i=1}^{n} |w_i|)^q \leq \sum_{i=1}^{n} |w_i|^q \leq (\sum_{i=1}^{n} |w_i|)^q. \end{cases} \tag{2.2}$$

Lemma 2.2. *[29] If $x, y \in \mathbb{R}$ and $a, b, c \geq 0$ hold, the following relationship exists:*

$$|x|^b |y|^c \leq \frac{ab}{b+c} |x|^{b+c} + \frac{c a^{-\frac{b}{c}}}{b+c} |y|^{b+c}. \tag{2.3}$$

Lemma 2.3. *[30] Consider the following dynamic system:*

$$\begin{cases} \dot{x}(t) = f(t, x(t)), \\ x(0) = x_0, \end{cases} \tag{2.4}$$

where $x(t)$=$[x_1(t), x_2(t), \cdots, x_n(t)]^T \in \mathbb{R}^n$*, and* $f(t, x(t)) \in \mathbb{R}^n \times \mathbb{R}^+ \rightarrow \mathbb{R}^n$ *is a nonlinear function with the initial value* $f(0) = 0$*. It is assumed that there is a function* $V(x_i(t), t)$ *satisfying* $V(x_i(t), t) \geq 0$ *and* $V(x) = 0 \Leftrightarrow x = 0$*. The derivative of this function,* $\dot{V}(x_i(t), t)$*, satisfies:*

$$\dot{V}(x_i(t), t) \leq -aV^c(x_i(t), t) - bV^d(x_i(t), t) + \tau, \tag{2.5}$$

where $a, b, \tau > 0$*,* $c \in (0, 1)$*, and* $d \in (1, +\infty)$*, it can be concluded that system (2.4) is practically fixed-time stable, and there is a convergence time* $T(x_0)$ *satisfying:*

$$T(x_0) \leq T_{max} = \frac{1}{a\kappa(1-b)} + \frac{1}{c\kappa(1-d)}, \tag{2.6}$$

where $\kappa \in (0, 1)$*. The residual error set of the system is*

$$\left\{ \lim_{t \to T(x_0)} x | V(x_i(t), t) \leq min \left\{ a^{-\frac{1}{c}} (\frac{\tau}{1-\kappa})^{\frac{1}{c}}, b^{-\frac{1}{d}} (\frac{\tau}{1-\kappa})^{\frac{1}{d}} \right\} \right\}. \tag{2.7}$$

Lemma 2.4. *[31] For a positive continuous function* $U(t)$ *with a bounded initial value* $U(0)$*, if its derivative satisfies.* $\dot{U}(t) \leq -sU(t) + k$*, where* s *and* k *are positive constants, then* $U(t)$ *satisfies the following:*

$$U(t) \leq -e^{st} U(0) + \frac{s}{k}(1 - e^{-st}). \tag{2.8}$$

2.3. Optimal control theory

The dynamic equation model for a nonlinear system is as follows [14]:

$$\dot{\chi}(t) = u(\chi(t)) + F(\chi(t)), \tag{2.9}$$

where $\chi \in \Omega \in \mathbb{R}^N$ denotes the state of the system, $u(\chi) \in \mathbb{R}^N$ represents the control input of the system, and $F(\chi) \in \mathbb{R}^N \times \mathbb{R}^+ \rightarrow \mathbb{R}^N$ denotes the nonlinear function of the system. The optimal control design entails the synthesis of a control input $u(\chi)$ over the set Ω (which contains the origin) to achieve the optimization of a predefined performance index. The objective function is designed as follows:

$$J(\chi) = \int_t^{+\infty} R(\chi(\tau), u(\chi(\tau))) d\tau. \tag{2.10}$$

The integrand is designed as $R(\chi(\tau), u(\chi(\tau))) = Q(\chi) + G(u(\chi))$, where the function $Q(\chi)$ is positive definite and satisfies $Q(0) = 0$ if and only if $\chi = 0$. $G(u(\chi))$ is a continuous function with respect to the control input $u(\chi)$ defined on the set Ω.

To solve the optimization problem, the first step is to minimize the objective function. Next, the minimized objective function is used to construct the Hamilton–Jacobi–Bellman (HJB) equation. Finally, the optimal control input $u^x(\chi)$ is derived from the HJB equation. The main process is as follows.

Minimize the following objective function:

$$J^x(\chi) = \min_{\chi \in \Omega} \int_t^{+\infty} R(\chi(\tau), u^x(\chi(\tau)))d\chi. \tag{2.11}$$

Construct the HJB equation using $J^x(\chi)$ as follows:

$$H^x(\chi, u^x, J^x) = R(\chi(\tau), u^x(\chi(\tau))) + \frac{\partial J^x}{\partial S(\chi)} \dot{S}(\chi) = 0, \tag{2.12}$$

where $S(\chi)$ is a continuous function defined on Ω, whose formulation involves $Q(\chi)$. The optimal control input $u_x(\chi_0)$ is then derived by considering the partial derivative of the constructed HJB equation with respect to u_x and setting it to zero ($\frac{\partial H_x(\chi, u_x, J_x)}{\partial u_x} = 0$).

2.4. The neural network state observer design

The design of the NN state observer is mainly aimed at solving the problem of unmeasurable states in the system (2.1). NNs are used to approximate unknown functions, and the approximate results are applied to the design of the observer. To realize the observer design, the basic principles of NNs are obtained from the literature [32] and applied to the observer design.

Lemma 2.5. *If $f(x)$ is a continuous function defined on a compact set Ω, then there exist ω^*, $\psi(x)$, and $\epsilon(x)$ such that the following equality holds:*

$$f(x) = \omega^{*T} \psi(x) + \epsilon(x), \tag{2.13}$$

where $\epsilon(x) > 0$ denotes the approximation error. There is a $\bar{\epsilon} > 0$ such that $\|\epsilon(x)\| \leq \bar{\epsilon}$ holds. Morover, $\omega^* = [\omega_1, \omega_2, \cdots, \omega_s] \in \mathbb{R}^s$ is the optimal weight vector and $\|\omega^*\| \leq \bar{\omega}, \bar{\omega} \geq 0$. $\psi(x) = [\psi_1(x), \cdots, \psi_s(x)]^T \in \mathbb{R}^s$ is the activation function vector, which is defined as

$$\psi_i(x) = exp\Big[\frac{-(x-\mu_i)^T(x-\mu_i)}{\eta_i^2}\Big], \tag{2.14}$$

where μ_i and η_i are the center and width of the activation function, respectively.

Since there exist unknown nonlinear functions $f_{i,j}(\bar{x}_{i,j})$ $(i = 1, 2, \cdots, N,\ j = 1, 2, \cdots, n)$ exist in the system (2.1), we can approximate these nonlinear functions using the approximation principle in Lemma 2.5, and the results are as follows:

$$f_{i,j}(\bar{x}_{i,j}) = \omega_{fi,j}^{*T} \psi_{fi,j}(\bar{x}_i) + \epsilon_{fi,j}, \tag{2.15}$$

where $\omega_{fi,j}^*$, $\psi_{fi,j}(\bar{x}_i)$, and $\epsilon_{fi,j}$ represent the weight vector, activation function, and approximation residual, respectively.

Rewrite the system (2.1) as follows:

$$\begin{cases} \dot{x}_{i,j}(t) = x_{i,j+1}(t) + \omega_{fi,j}^{*T} \psi_{fi,j}(\bar{x}_i) + \epsilon_{fi,j}, \\ \dot{x}_{i,n}(t) = u_i(t) + \omega_{fi,n}^{*T} \psi_{fi,n}(\bar{x}_i) + \epsilon_{fi,n}, \\ y_i(t) = x_{i,1}(t), \\ i = 1, 2, \cdots, N, j = 1, 2, \cdots, n-1. \end{cases} \tag{2.16}$$

Construct a NN observer as follows:

$$\begin{cases}\dot{\hat{x}}_{i,j}(t) = \hat{x}_{i,j+1}(t) + \hat{\omega}^T_{fi,j}\psi_{fi,j}(\bar{x}_i),\\ \dot{\hat{x}}_{i,n}(t) = u_i(t) + \hat{\omega}^T_{fi,n}\psi_{fi,n}(\bar{x}_i),\\ \hat{y}_i(t) = \hat{x}_{i,1}(t),\\ i = 1, 2, \cdots, N, j = 1, 2, \cdots, n-1,\end{cases} \tag{2.17}$$

where $\hat{x}_{i,j}(t)$ and $\hat{\omega}_{fi,j}$ are the observed values of $x_{i,j}(t)$ and $\omega^*_{fi,j}$, respectively. Further, rewrite Eqs (2.16) and (2.17) in the following forms:

$$\begin{cases}\dot{H}_i = A_iH_i + B_iy_i + \sum_{j=1}^n Q_{i,j}\omega^{*T}_{fi,j}\psi_{fi,j}(\bar{x}_i) + K_nu_i + \epsilon_i,\\ y_i = Cx_{i,1},\end{cases} \tag{2.18}$$

$$\begin{cases}\dot{\hat{H}}_i = A_i\hat{H}_i + B_i\hat{y}_i + \sum_{j=1}^n Q_{i,j}\hat{\omega}^T_{fi,j}\psi_{fi,j}(\bar{x}_i) + K_nu_i,\\ \hat{y}_i = C\hat{x}_{i,1}.\end{cases} \tag{2.19}$$

For the system (2.18), where H_i=$[x_{i,1}, x_{i,2}, \cdots, x_{i,n}]^T$, C=$[1, 0, \cdots, 0]_{n\times 1}$, ϵ_i=$[\epsilon_{i,1}, \cdots, \epsilon_{i,n}]^T$, $K_n = [0, \cdots, 1]^T_{n\times 1}$, $B_i = [w_{i,1}, w_{i,2}, \cdots, w_{i,n}]^T$, and $Q_{i,j} = [\underbrace{0, \ldots, 1}_{j}, \ldots, 0]^{\mathrm{T}}_{n\times 1}$, define the matrix A_i as follows:

$$A_i = \begin{bmatrix} -w_{i,1} & & & \\ \vdots & & I_{n-1} & \\ -w_{i,n} & 0 & \cdots & 0\end{bmatrix}_{n\times n}.$$

The system (2.19) is the observation result of the system (2.18), where $\hat{H}_i = [\hat{x}_{i,1}, \hat{x}_{i,2}, \cdots, \hat{x}_{i,n}]^T$. Define the observation error e_i as follows:

$$e_i = [e_{i,1}, \cdots, e_{i,n}]^T = H_i - \hat{H}_i = \bar{x}_{i,n} - \hat{\bar{x}}_{i,n}, \tag{2.20}$$

where $\hat{\bar{x}}_{i,n} = [\hat{x}_{i,1}, \cdots, \hat{x}_{i,n}]^T$.

By combining Eqs (2.18)–(2.20), we can obtain:

$$\dot{e}_i = A_ie_i + \sum_{j=1}^{n} Q_{i,j}\tilde{\omega}_{fi,j}\psi_{fi,j}(\bar{x}_{i,j}) + \epsilon_i, \tag{2.21}$$

where $\tilde{\omega}_{fi,j} = \hat{\omega}_{fi,j} - \omega^*_{fi,j}$. Define $\dot{\hat{\omega}}_{fi,j}$ as follows:

$$\dot{\hat{\omega}}_{fi,j} = -\theta_{i,j}\psi^T_{fi,j}(\bar{x}_{i,j})\psi_{fi,j}(\bar{x}_{i,j})\hat{\omega}_{fi,j}. \tag{2.22}$$

Heres, $\theta_{i,j}(i = 1, \cdots, N, j = 1, \cdots, n)$ is a positive constant. By selecting an appropriate vector B_i, the matrix A_i can be made a strictly Hurwitz matrix. Choosing a positive definite matrix $Q_i = Q_i^T > 0$ and a symmetric matrix $P_i = P_i^T$, we can obtain the following Riccati equation:

$$A_i^TP_i + P_iA_i = -2Q_i. \tag{2.23}$$

3. Convergence analysis and main results

3.1. Boundedness proof

In this section, we prove that the observation error is bounded on the basis of Lemma 2.4, discuss its impact on subsequent reinforcement learning algorithms, and elaborate on how to reduce such error.

Theorem 3.1. *Assuming that the update law (2.22) holds, it can be concluded that both the observation error e_i and $\tilde{\omega}_{fi,j}$ are bounded. Moreover, by selecting appropriate observation gain coefficients $w_{i,j}$, the observation error e_i can converge to the desired precision level.*

Construct the following Lyapunov function:

$$V(t) = \frac{1}{2}\sum_{i=1}^{N} e_i^T P_i e_i + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{n} \tilde{\omega}_{fi,j}^T \tilde{\omega}_{fi,j}, \tag{3.1}$$

according to Eq (2.23). Taking the derivative of V yields:

$$\begin{aligned}\dot{V}(t) \leq& \sum_{i=1}^{N} e_i^T P_i (Ae_i + \sum_{j=1}^{n} Q_{i,j}\tilde{\omega}_{fi,j}\psi_{fi,j}^T(\bar{x}_{i,j}) + \epsilon_i) \\ &+ \sum_{i=1}^{N} \tilde{\omega}_{fi,j}^T \theta_{i,j}\psi_{fi,j}^T(\bar{x}_{i,j})\psi_{fi,j}(\bar{x}_{i,j})\hat{\omega}_{fi,j} \\ \leq& \sum_{i=1}^{N} \frac{1}{2} e_i^T (A_i^T P_i + P_i A_i) e_i \\ &+ \sum_{i=1}^{N} e_i^T P_i (\sum_{j=1} Q_{i,j}\tilde{\omega}_{fi,j}\psi_{fi,j}^T(\bar{x}_{i,j}) + \epsilon_i) \\ &+ \sum_{i=1}^{N}\sum_{j=1}^{n} \tilde{\omega}_{fi,j}^T \theta_{i,j}\psi_{fi,j}^T(\bar{x}_{i,j})\psi_{fi,j}(\bar{x}_{i,j})\hat{\omega}_{fi,j}.\end{aligned} \tag{3.2}$$

According to Young's inequality, the following formula can be obtained

$$e_i^T P_i \epsilon_i \leq \|P_i\|^2 \bar{\epsilon}_i^{\,2} + \|e_i\|^2, \tag{3.3}$$

$$0 < \psi_{fi,j}^T(\bar{x}_{i,j})\psi_{fi,j}(\bar{x}_{i,j}) < l_i, (0 < l_i < 1), \tag{3.4}$$

$$\tilde{\omega}_{fi,j}^T \psi_{fi,j}^T(\bar{x}_{i,j})\psi_{fi,j}(\bar{x}_{i,j})\hat{\omega}_{fi,j} \leq \frac{1}{2} l_i \omega_{fi,j}^{*T}\omega_{fi,j}^{*} - \frac{1}{2} l_i \hat{\omega}_{fi,j}^T \hat{\omega}_{fi,j} - \frac{1}{2} l_i \tilde{\omega}_{fi,j}^T \tilde{\omega}_{fi,j}, \tag{3.5}$$

$$e_i^T P_i (\sum_{j=1}^{n} Q_{i,j}\tilde{\omega}_{fi,j}\psi_{fi,j}^T(\bar{x}_{i,j})) \leq \frac{1}{q}\|P_i\|^2 l_i \sum_{j=1}^{n} \tilde{\omega}_{fi,j}^T \tilde{\omega}_{fi,j} + q\|e_i\|^2. \tag{3.6}$$

Substituting (2.22) and (3.3)–(3.6) into (3.2) yields:

$$\begin{aligned}\dot{V}(t) \leq& -\sum_{i=1}^{N} \lambda_{min}(Q_i)\|e_i\|^2 + \sum_{i=1}^{N} \|e_i\|^2 + \sum_{i=1}^{N} q\|e_i\|^2 - \sum_{i=1}^{N}\sum_{j=1}^{n} \frac{1}{2} l_i \theta_{i,j} \tilde{\omega}_{fi,j}^T \tilde{\omega}_{fi,j} \\ &+ \sum_{i=1}^{N} \frac{1}{q}\|P_i\|^2 l_i \sum_{j=1}^{n} \tilde{\omega}_{fi,j}^T \tilde{\omega}_{fi,j} + \sum_{i=1}^{N}\sum_{j=1}^{n} \frac{1}{2} l_i \theta_{i,j} \omega_{fi,j}^{*T}\omega_{fi,j}^{*} + \sum_{i=1}^{N} \|P_i\|^2 \bar{\epsilon}_i^2.\end{aligned} \tag{3.7}$$

Further rearrangement yields

$$\dot{V}(t) \leq -\sum_{i=1}^{N} \lambda_0 \|e_i\|^2 - \sum_{i=1}^{N}\sum_{j=1}^{n} \mu_0 \tilde{\omega}_{fi,j}^T \tilde{\omega}_{fi,j} + M_0, \tag{3.8}$$

where we let $\lambda_0 = \min\{\lambda_{i,0} = \lambda_{\min}(Q_i) - 1 - q\}$, and let $\mu_0 = \frac{1}{2}l_0\theta_0 - \frac{1}{q}\|P_0\|^2 l_0$ be the minimum value of $\frac{1}{2}l_i\theta_{i,j} - \frac{1}{q}\|P_i\|^2 l_i$ for $i = 1, 2, \cdots, N$. Given $\sum_{i=1}^{N}\sum_{j=1}^{n} \frac{1}{2} l_i \omega_{fi,j}^{*T}\omega_{fi,j}^{*} + \sum_{i=1}^{N} \|P_i\|^2 \bar{\epsilon}_i^2 \leq M_0$, take $\tau_0 = \min\{\lambda_0, \mu_0\}$. To summarize, Eq (3.8) can be rewritten as

$$\dot{V}(t) \leq -\tau_0 V(t) + M_0. \tag{3.9}$$

According to Lemma 2.4, we have $V(t) \leq -e^{-\tau_0 t}V(0) + \frac{M_0}{\tau_0}(1 - e^{-\tau_0 t})$. Thus, $V(t)$ is bounded and Theorem 3.1 is proved.

Remark 3.1. *From (3.9), we can see that the magnitude of M_0 affects the size of the error. To minimize the error as much as possible, we need to reduce the value of $\sum_{i=1}^{N}\sum_{j=1}^{n} \frac{1}{2} l_i \omega_{fi,j}^{*T}\omega_{fi,j}^{*} + \sum_{i=1}^{N} \|P_i\|^2 \bar{\epsilon}_i^2$. This can be achieved by reducing the number of network layers and neurons to decrease the number of parameters. Thus, we adopt a one-dimensional vector $\omega_{fi,j} \in \mathbb{R}^{n\times 1}$. Additionally, we reduce the value of l_i (referring to the results in [29], where $l_i \leq 1$, and minimize the error $\|\epsilon_{fi,j}\|$ by setting*

$$\sigma_{i,j}^{*} = \arg \min_{\omega_{fi,j}^{*} \in \Omega_i} \left\{ \sup_{\bar{x}_i \in S_i} \|\omega_{fi,j}^{*T} \psi_{fi,j}(\bar{x}_{i,j}) - f_{i,j}(\bar{x}_{i,j})\| \right\},$$

Ω_i is the set containing all $\omega_{fi,j}$, and S_i is the set containing all $\bar{x}_{i,j}$. Let$\|\epsilon_{fi,j}\| \leq \bar{\epsilon}_i \leq \sigma_{i,j}^{*}$

Remark 3.2. *As indicated in Section 2.3, the core objective of the RL algorithm is to obtain the optimal control input, which essentially refers to the minimum control input. When the control input reaches the minimum value, resource consumption is also minimized, and the control input achieves the most stable state at this time. To achieve this goal, we need to find a sufficiently small cost function, so we usually select a cost function that includes an error term, and its minimization is relatively straightforward. Herein, we only prove that the error is bounded, meaning there is further room to minimize the cost function. We can reduce the error by decreasing the value of M_0 and increasing the value of τ_0.*

3.2. Selection of objective function

Define the following coordinate variables:

$$z_{i,1} = \sum_{k=1}^{n} a_{i,k}(\hat{y}_i - \hat{y}_k) + d_i(\hat{y}_i - y_d)(i = 1, \cdots, N), \tag{3.10}$$

$$g_{i,1} = \hat{x}_{i,1} - y_d, \tag{3.11}$$

$$g_{i,j} = \hat{x}_{i,j} - \delta_{i,j-1}^{x}(j = 2, \cdots, n). \tag{3.12}$$

Here, $\delta_{i,j}^{x}$ denotes the virtual control input, with $\delta_{i,0}^{x} = y_d$ and $\hat{\delta}_{i,n}^{x} = \hat{u}_i^{x}$, and $\hat{\delta}_{i,j}^{x}$ is the observed value of the virtual control input.

Differentiating Eqs (3.10)–(3.12) also yields:

$$\dot{z}_{i,1} = \pi_{i,1}(\hat{x}_{i,2} + \hat{\omega}_{fi,1}^T \psi_{fi,1} - \dot{y}_d) - \sum_{i=1}^{N} a_{i,k}\dot{g}_{i,1}, \tag{3.13}$$

$$\dot{g}_{i,j} = \hat{x}_{i,j+1} + \hat{\omega}_{fi,j}^T \psi_{fi,j} - \dot{\delta}_{i,j-1}^x. \tag{3.14}$$

According to the optimal control theory in Subsection 2.3, to obtain the optimal control input, we need to minimize the objective function and construct the HJB equation. Therefore, the selection of the objective function is crucial. The optimal objective function will be constructed below, based on the coordinate changes in Eqs (3.10)–(3.12).

Remark 3.3. *In this process, the backstepping method is employed to solve for the optimal control input $\delta_{i,j}^x$. The core idea of the backstepping method lies in backtracking the problem on the basis of the assumed validity of the result. Since we need to prove the convergence of Eqs (3.11) and (3.12), the backstepping method allows us to derive $\hat{x}_{i,j} \to \delta_{i,j-1}^x$. Therefore, during the derivation of $\delta_{i,j}^x$, $\delta_{i,j-1}^x$ is used to replace $\hat{x}_{i,j}$ for calculation.*

(1) Based on $g_{i,j}$ and $\delta_{i,j}$, the objective function $J^1(g_{i,j}, \delta_{i,j})$ is as follows:

$$J^1(g_{i,j}, \delta_{i,j}) = \int_t^{+\infty} (g_{i,j}^2(\chi) + \delta_{i,j}^2(\chi))d\chi. \tag{3.15}$$

Minimizing Eq (3.15) yields:

$$J^{1x}(g_{i,j}, \delta_{i,j}) = \min_{\delta_{i,j}\in\Omega_x} \int_t^{+\infty} (g_{i,j}^2(\chi) + \delta_{i,j}^{2x}(\chi))d\chi. \tag{3.16}$$

Here, the set Ω_x contains the origin and all real control inputs $\delta_{i,j}$ $(i = 1, \cdots, N,\ j = 1, \cdots, n)$. Further, the HJB equation is constructed as follows:

$$H^{1x}(g_{i,j}, \delta_{i,j}, \frac{\partial J^{1x}(g_{i,j}, \delta_{i,j})}{\partial g_{i,j}}) = g_{i,j}^2 + \delta_{i,j}^{x2} + \frac{\partial J^{1x}(g_{i,j}, \delta_{i,j})}{\partial g_{i,j}}\dot{g}_{i,j} = 0. \tag{3.17}$$

Taking the partial derivative of Eq (3.17) with respect to the control input, one has

$$\delta_{i,j}^x = -\frac{1}{2}\pi_{i,j}\frac{\partial J^{1x}(g_{i,j}, \delta_{i,j})}{\partial g_{i,j}}, \tag{3.18}$$

where $\pi_{i,j} = \begin{cases} \sum_{j=1}^{n} a_{i,j} + d_i,\ j = 1, \\ 1,\ j = 2, 3, \cdots, n. \end{cases}$

(2) Based on $\dot{g}_{i,j}$ and $g_{i,j}$, the objective function $J^2(g_{i,j}, \dot{g}_{i,j})$ is as follows:

$$J^2(g_{i,j}, \dot{g}_{i,j}) = \int_t^{+\infty} (g_{i,j}^2(\chi) + \dot{g}_{i,j}^2(\chi))d\chi. \tag{3.19}$$

Minimizing the objective function yields

$$J^{2x}(g_{i,j}, \dot{g}_{i,j}) = \min_{\delta_{i,j}\in\Omega_x} \int_t^{+\infty} (g_{i,j}^2(\chi) + \dot{g}_{i,j}^2(\chi))d\chi. \tag{3.20}$$

Constructing the HJB equation yields

$$H^{2x}(g_{i,j}, \delta_{i,j}, \frac{\partial J^{2x}(g_{i,j}, \dot{g}_{i,j})}{\partial g_{i,j}}) = g_{i,j}^2 + \dot{g}_{i,j}^2 + \frac{\partial J^{2x}(g_{i,j}, \dot{g}_{i,j})}{\partial g_{i,j}} \dot{g}_{i,j} = 0. \tag{3.21}$$

Taking the partial derivative of Eq (3.21) with respect to the control input and taking the optimal control input value $\delta_{i,j}^x$ at its minimum point, we obtain:

$$\delta_{i,j}^x = -\frac{1}{2}\pi_{i,j} \frac{\partial J^{1x}(g_{i,j}, \dot{g}_{i,j})}{\partial g_{i,j}} - \hat{\omega}_{fi,j}^T \psi_{fi,j} + \dot{\delta}_{i,j-1}^x. \tag{3.22}$$

Remark 3.4. *The core process of deriving the optimal control input is to minimize the objective function, where the minimization mainly refers to minimizing the control input $\delta_{i,j}$. Therefore, the objective function must contain $\delta_{i,j}$. Unlike the first cost function, which directly minimizes $\delta_{i,j}$, the second cost function minimizes $\dot{\hat{g}}_{i,j}$, which includes $\delta_{i,j}$. As a result, the optimal control input derived from the second objective function exhibits a more complex structure compared with that obtained from the first one. Nevertheless, the second formulation incorporates a broader set of factors, leading to a control law that better aligns with the design expectations. To validate this claim, the following analysis will evaluate both objective functions in terms of non-negativity, boundedness, and stability.*

For the first objective function, its non-negativity is obvious, so we consider its stability and boundedness. We construct a Lyapunov function related to the objective function and use the Lyapunov stability theory for the proof. The selected Lyapunov function is as follows:

$$V_1 = J_{i,j}^1(g_{i,j}, \delta_{i,j}). \tag{3.23}$$

It is always greater than 0. Taking its derivative with respect to time t yields

$$\dot{V}_1 = \frac{\partial J_{i,j}^1(g_{i,j}, \delta_{i,j})}{\partial t} = \frac{\partial J_{i,j}^1(g_{i,j}, \delta_{i,j})}{\partial g_{i,j}} \frac{\partial g_{i,j}}{\partial t} = \frac{\partial J_{i,j}^1}{\partial g_{i,j}} (\hat{\omega}_{fi,g}^T \psi_{fi,j} - \dot{\hat{\delta}}_{i,j}^x) - \frac{1}{2}(\frac{\partial J_{i,j}^1}{\partial g_{i,j}})^2. \tag{3.24}$$

We can only determine that $-\frac{1}{2}\left(\frac{\partial J_{i,j}^1}{\partial g_{i,j}}\right)^2 \leq 0$ holds, but cannot judge the sign of $\frac{\partial J_{i,j}^1}{\partial g_{i,j}}\left(\hat{\omega}_{fi,g}^T \psi_{fi,j} - \dot{\delta}_{i,j-1}^x\right)$. Therefore, the stability of the first objective function $J_{i,j}^1(g_{i,j}, \delta_{i,j})$ cannot be determined. Similarly, its boundedness cannot be determined.

For the second objective function, its non-negativity is also obvious due to the presence of squared terms and the characteristics of the integral interval. Similar to the first objective function, we align construct a Lyapunov function to discuss its stability and boundedness, as follows:

$$V_2 = J^2(g_{i,j}, \delta_{i,j}), \tag{3.25}$$

$V_2 \geq 0$ holds. Taking its derivative yields:

$$\dot{V}_2 = \frac{\partial J^2(g_{i,j}, \delta_{i,j})}{\partial g_{i,j}} \frac{\partial g_{i,j}}{\partial t} = -\frac{1}{2}(\frac{\partial J_{i,j}^2}{\partial g_{i,j}})^2 \leq 0. \tag{3.26}$$

From Eq (3.26) and Lyapunov stability theory, it can be concluded that the second objective function $J_{i,j}^2(g_{i,j}, \delta_{i,j})$ tends to be stable. The boundedness is also guaranteed.

Remark 3.5. *Compared with the first objective function $J^1(g_{i,j}, \delta_{i,j})$, the second objective function $J^2(g_{i,j}, \delta_{i,j})$ has its non-negativity, boundedness, and stability all guaranteed. In the process of deriving the optimal control input $\delta_{i,j}$, the core is to minimize the objective function to achieve the minimization of the control input; however, reducing the error ($g_{i,j}$) is also a crucial task. It can be observed that the construction of the objective function is inherently linked to both the tracking error and the control input. This is primarily because minimizing the objective function requires concurrently reducing the error $g_{i,j}$ and optimizing the control effort $\delta_{i,j}$, which together constitute the desired design objectives. Therefore, the objective function must include either the error and the control input directly or terms containing the control input. Meanwhile, we need to consider the entire optimization process, and the existence of the integral sign avoids achieving only instantaneous optimality.*

Remark 3.6. *The first objective function has been presented in many works such as [7, 24, 25, 31]. Moreover, it remains the most frequently adopted form of objective function in the related literature. In contrast to previous studies, this work further analyzes the objective function from the perspectives of non-negativity, boundedness, and stability. However, its stability cannot be determined, so we propose a second objective function with stronger stability. Its advantages lie in enhancing the robustness of the system; moreover, a stable objective function can make the optimization process smoother, avoid falling into local optimality, and shorten the convergence time.*

4. Reinforcement learning based on objective functions

In the previous section, we selected an objective function that is suitable for the feedback system in this paper. In this section, we will derive specific optimal control inputs using this objective function. We will divide the process into n steps, thereby obtaining n optimal control inputs $\delta^x_{i,j}$.

Step (1). The selected cost function is

$$J(z_{i,1}, \dot{g}_{i,1}) = \int_t^{+\infty} \left(z^2_{i,1}(\chi) + \dot{g}^2_{i,1}(\chi)\right) d\chi. \tag{4.1}$$

Minimizing the objective function yields $J^x(g_{i,1}, \delta_{i,1}) = \min\limits_{\delta_{i,j}\in\Omega_x} \int_t^{+\infty} \left(z^2_{i,1}(\chi) + \dot{g}^2_{i,1}\right) d\chi$.

According to the minimized objective function, the resulting HJB equation is:

$$H^x(z_{i,1}, \dot{g}_{i,1}) = z^2_{i,1} + \dot{g}^2_{i,1} + \frac{\partial J^x(z_{i,1}, \dot{g}_{i,1})}{\partial z_{i,1}} \dot{z}_{i,1} = 0. \tag{4.2}$$

Substituting Eqs (3.10) and (3.12), using the backstepping method where $x_{i,j+1} \to \delta^x_{i,j}$, and taking the value of $\delta_{i,1}$ at the point where $\frac{\partial H^x(z_{i,1},\dot{g}_{i,1})}{\partial \delta_{i,1}} = 0$, we obtain:

$$2(\delta^x_{i,1} + \hat{\omega}^T_{fi,1}\psi_{fi,1} - \dot{y}_d) + \pi_{i,1}\frac{\partial J^x(z_{i,1}, \dot{g}_{i,1})}{\partial z_{i,1}} = 0, \tag{4.3}$$

$$\delta^x_{i,1} = -\frac{1}{2}\pi_{i,j}\frac{\partial J^x(z_{i,1}, \dot{g}_{i,1})}{\partial z_{i,1}} - \hat{\omega}^T_{fi,1}\psi_{fi,1} + \dot{y}_d. \tag{4.4}$$

Next, we will reconstruct $\frac{\partial J^x(z_{i,1},\dot{g}_{i,1})}{\partial z_{i,1}}$ as follows:

$$\frac{\partial J^x(z_{i,1}, \dot{g}_{i,1})}{\partial z_{i,1}} = \frac{2\eta_{i,1}}{\pi_{i,1}}(g^T_{i,1}g_{i,1})g_{i,1} + \frac{2\nu_{i,1}}{\pi_{i,1}}(g^T_{i,1}g_{i,1})^{p-1}g_{i,1} + 4g_{i,1} + \frac{2}{\pi_{i,1}}h_{i,1}(x_{i,1}) + \frac{1}{\pi_{i,1}}J^{x_1}, \tag{4.5}$$

$$J^{x_1} = -2\eta_{i,1}(g_{i,1}^T g_{i,1})g_{i,1} - 2\nu_{i,1}(g_{i,1}^T g_{i,1})^{p-1} g_{i,1} - 4g_{i,1} - 2h_{i,1}(x_{i,1}) + \pi_{i,1}\frac{\partial J^x(z_{i,1}, \dot{g}_{i,1})}{\partial z_{i,1}}. \tag{4.6}$$

Substituting Eqs (4.5) and (4.6) into Eq (4.4), we have

$$\delta_{i,1}^x = -\eta_{i,1}(g_{i,1}^T g_{i,1})g_{i,1} - \nu_{i,1}(g_{i,1}^T g_{i,1})^{p-1} g_{i,1} - 2g_{i,1} - h_{i,1}(x_{i,1}) - \frac{1}{2}J^{x_1} - \hat{\omega}_{fi,1}^T \psi_{fi,1} + \dot{y}_d, \tag{4.7}$$

where $\eta_{i,1}$ and $\nu_{i,1}$ are both positive constants ($i= 1, 2, \cdots, N$) and $P \in (0, 1)$.

Remark 4.1. *Note that $h_{i,1}(x_{i,1})$ is an introduced unknown nonlinear function. Its introduction is to ensure the unknownness of J^{x_1} so that NNs can be used to approximate J^{x_1}, thereby handling the partial derivative term $\frac{\partial J^x(z_{i,1}, \dot{g}_{i,1})}{\partial z_{i,1}}$. In the process of approximating J^{x_1}, two weight vectors $\hat{\omega}_{ai,1}$ and $\hat{\omega}_{ci,1}$ will be obtained, which are crucial for us to achieve the FxT consensus of the system.*

For the unknown nonlinear functions $h_{i,1}(x_{i,1})$ and J^{x_1}, by combining them NNS to approximate them, we can obtain

$$h_{i,1}(x_{i,1}) = \omega_{h_i,1}^{*T}\psi_{h_i,1}(x_{i,1}) + \epsilon_{h_i,1}, \tag{4.8}$$

$$J^{x_1} = \omega_{J_1}^{*T}\psi_{J_1} + \epsilon_{J_1}, \tag{4.9}$$

where $\omega_{h_i,1}^{*} \in \mathbb{R}^{m_{11}\times n}$, $\omega_{J_1}^{*T} \in \mathbb{R}^{m_{12}\times n}$, $\psi_{h_i,1} = [\psi_{h_i,11}, \psi_{h_i,12}, \cdots, \psi_{h_i,1m_{11}}]^T$,and $\psi_{J_1}=[\psi_{J_1,1} \cdots, \psi_{J_1,m_{12}}]$. Substituting them into Eqs (4.6) and (4.7), we obtain

$$\begin{aligned}\delta_{i,1}^x =& -\eta_{i,1}(g_{i,1}^T g_{i,1})g_{i,1} - \nu_{i,1}(g_{i,1}^T g_{i,1})^{p-1} g_{i,1} - 2g_{i,1} - \hat{\omega}_{fi,1}^T \psi_{fi,1} + \dot{y}_d \\ & - \omega_{h_i,1}^{*T}\psi_{h_i,1}(x_{i,1}) - \epsilon_{h_i,1} - \frac{1}{2}\omega_{J_1}^{*T}\psi_{J_1} - \frac{1}{2}\epsilon_{J_1},\end{aligned} \tag{4.10}$$

$$\begin{aligned}\frac{\partial J^x(z_{i,1}, \dot{g}_{i,1})}{\partial z_{i,1}} =& \frac{2\eta_{i,1}}{\pi_{i,1}}(g_{i,1}^T g_{i,1})g_{i,1} + 4g_{i,1} + \frac{2\nu_{i,1}}{\pi_{i,1}}(g_{i,1}^T g_{i,1})^{p-1} g_{i,1} \\ & + \frac{2}{\pi_{i,1}}\omega_{h_i,1}^{*T}\psi_{h_i,1}(x_{i,1}) + \frac{2}{\pi_{i,1}}\epsilon_{h_i,1} + \frac{1}{\pi_{i,1}}\omega_{J_1}^{*T}\psi_{J_1} + \frac{1}{\pi_{i,1}}\epsilon_{J_1}.\end{aligned} \tag{4.11}$$

Next, the control performance is evaluated on the basis of the following critic network

$$\begin{aligned}\frac{\partial \hat{J}^x(z_{i,1}, \dot{g}_{i,1})}{\partial z_{i,1}} =& \frac{2\eta_{i,1}}{\pi_{i,1}}(g_{i,1}^T g_{i,1})g_{i,1} + 4g_{i,1} + \frac{2\nu_{i,1}}{\pi_{i,1}}(g_{i,1}^T g_{i,1})^{p-1} g_{i,1} \\ & + \frac{2}{\pi_{i,1}}\omega_{h_i,1}^{*T}\psi_{h_i,1}(x_{i,1}) + \frac{2}{\pi_{i,1}}\epsilon_{h_i,1} + \frac{1}{\pi_{i,1}}(\hat{\omega}_{ci,1}^T\psi_{i,1}),\end{aligned} \tag{4.12}$$

where $\hat{\omega}_{ci,1}$ denotes the weight vector of the critic network, and its update law satisfies the following equation:

$$\dot{\hat{\omega}}_{ci,1} = -\beta_{ci,1}\psi_{i,1}\psi_{i,1}^T\hat{\omega}_{ci,1}. \tag{4.13}$$

Next, we construct an actor network based on NN technology:

$$\begin{aligned}\hat{\delta}_{i,1}^x =& -\eta_{i,1}(g_{i,1}^T g_{i,1})g_{i,1} - \nu_{i,1}(g_{i,1}^T g_{i,1})^{p-1} g_{i,1} - 2g_{i,1} - \hat{\omega}_{fi,1}^T \psi_{fi,1} \\ & + \dot{y}_d - \omega_{h_i,1}^{*T}\psi_{h_i,1}(x_{i,1}) - \epsilon_{h_i,1} - \frac{1}{2}(\hat{\omega}_{ai,1}^T\psi_{i,1}),\end{aligned} \tag{4.14}$$

where $\hat{\omega}_{ai,1}$ is the weight vector of the actor network, and its weight update law satisfies the following equation:

$$\dot{\hat{\omega}}_{ai,1} = -\beta_{ai,1}\psi_{i,1}\psi_{i,1}^T(\hat{\omega}_{ai,1} - \frac{1}{2}\hat{\omega}_{ci,1}). \tag{4.15}$$

Remark 4.2. *Note that m_{12} denotes the structural dimension in the critic–actor network, i.e., the number of neurons. The structural dimensions of the critic's weight and the actor's weight obtained after evaluating and adjusting the optimal control input should have the same dimension, i.e., $\hat{\omega}_{ci,1} \in \mathbb{R}^{m_{12}\times n}$ and $\hat{\omega}_{ai,1} \in \mathbb{R}^{m_{12}\times n}$. It can be seen that when the critic–actor network evaluates and adjusts the optimal control input, it actually acts on J^{x_1}. Since J^{x_1} already covers all quantities of $\delta^x_{i,j}$, it is sufficient to evaluate and adjust J^{x_1}.We mainly use the critic–actor framework to evaluate and adjust the control input $\delta^x_{i,1}$. The critic framework evaluates the value of the control input, while the actor framework adapts its parameters online in response to this evaluation, thereby enabling the generated control input to achieve superior performance. Here, $\hat{\delta}^x_{i,1}$ is the control input obtained after evaluation and adjustment by the critic–actor network. Next, we use the Bellman residual to analyze the rationality of the obtained weights of the critic–actor network.*

Define the Bellman residual as

$$E_{i,1} = H^x(z_{i,1}, \dot{\hat{g}}_{i,1}) - H^x(z_{i,1}, \dot{g}_{i,1}) = H^x(z_{i,1}, \dot{\hat{g}}_{i,1}), \tag{4.16}$$

where $H^x(z_{i,1}, \dot{\hat{g}}_{i,1})$ is the HJB equation obtained by substituting $\frac{\partial \hat{J}^x(z_{i,1},\dot{g}_{i,1})}{\partial z_{i,1}}$ and $\hat{\delta}^x_{i,1}$ into $H^x(z_{i,1}, \dot{g}_{i,1})$. The selection of $\hat{\omega}_{ai,1}$ is such that, on the one hand, the Bellman residual $E_{i,1} = 0$, and on the other hand, $\frac{\partial E_{i,1}}{\partial \hat{\omega}_{ai,1}} = 0$.

$$\begin{aligned} E_{i,1} =& z_{i,1}^2 + (-\eta_{i,1}(g_{i,1}^T g_{i,1})g_{i,1} - \nu_{i,1}(g_{i,1}^T g_{i,1})^{p-1} g_{i,1} \\ & - \frac{1}{2} g_{i,1} - \omega^{*T}_{h_i,1}\psi_{h_i,1}(x_{i,1}) - \epsilon_{h_i,1} - \frac{1}{2}(\hat{\omega}^T_{ai,1}\psi_{i,1}))^2 \\ & + [\frac{2\eta_{i,1}}{\pi_{i,1}}(g_{i,1}^T g_{i,1})g_{i,1} + \frac{2\nu_{i,1}}{\pi_{i,1}}(g_{i,1}^T g_{i,1})^{p-1} g_{i,1} \\ & + 4g_{i,1} + \frac{2}{\pi_{i,1}}\omega^{*T}_{h_i,1}\psi_{h_i,1}(x_{i,1}) + \frac{2}{\pi_{i,1}}\epsilon_{h_i,1} \\ & + \frac{1}{\pi_{i,1}}(\hat{\omega}^T_{ci,1}\psi_{i,1})][-\eta_{i,1}(g_{i,1}^T g_{i,1})g_{i,1} \\ & - \nu_{i,1}(g_{i,1}^T g_{i,1})^{p-1} g_{i,1} - 2g_{i,1} - \hat{\omega}^T_{fi,1}\psi_{fi,1} + \dot{y}_d - \omega^{*T}_{h_i,1}\psi_{h_i,1}(x_{i,1}) \\ & - \epsilon_{h_i,1} - \frac{1}{2}(\hat{\omega}^T_{ai,1}\psi_{i,1})\pi_{i,1} + \hat{\omega}^T_{fi,1}\psi_{fi,1} - \sum_{i=1}^{N} a_{i,k}\hat{y}_k - d_i y_d]. \end{aligned} \tag{4.17}$$

The specific form of $\frac{\partial E_{i,1}}{\partial \hat{\omega}_{ai,1}} = 0$ is obtained as:

$$\frac{\partial E_{i,1}}{\partial \hat{\omega}_{ai,1}} = \frac{1}{2}\psi_{i,1}\psi_{i,1}^T(\hat{\omega}_{ai,1} - \hat{\omega}_{ci,1}) = 0. \tag{4.18}$$

Let $W_{i,1} = \hat{\omega}_{ai,1}\frac{\partial E_{i,1}}{\partial \hat{\omega}_{ai,1}} = 0$ and $\dot{\hat{\omega}}_{ai,1} = -\beta_{ai,1}\frac{\partial W_{i,1}}{\hat{\omega}_{ai,1}}$. The update law for the weight vector is obtained as:

$$\dot{\hat{\omega}}_{ai,1} = -\beta_{ai,1}\psi_{i,1}\psi_{i,1}^T(\hat{\omega}_{ai,1} - \frac{1}{2}\hat{\omega}_{ci,1}). \tag{4.19}$$

The weight update law for the actor network is derived on the basis of the Bellman residual, while the update law for the critic network is subsequently designed in dependency on the actor's weight adjustment. This coupling stems from the fact that within the critic–actor framework, the evaluation and refinement of the optimal control input inherently link the learning processes of both networks. The obtained weight vectors $\hat{\omega}_{ai,1}$ and $\hat{\omega}_{ci,1}$ are both derived from J^{x_1}, and these two weight vectors are approximately consistent. The update law for $\hat{\omega}_{ci,1}$ is

$$\dot{\hat{\omega}}_{ci,1} = -\beta_{ci,1}\psi_{i,1}\psi_{i,1}^T\hat{\omega}_{ci,1}. \tag{4.20}$$

Both $\beta_{ai,1}$ and $\beta_{ci,1}$ are constants greater than 0. To prove that the selected critic weights can converge to the actor weights, we construct a Lyapunov function as follows:

$$V_{e_{w_{i,1}}} = \frac{1}{2}Tr(w_{i,1}^T w_{i,1}), \tag{4.21}$$

where $w_{i,1} = \hat{\omega}_{ai,1} - \hat{\omega}_{ci,1}$. Taking the derivative of $V_{e_{w_{i,1}}}$, we obtain

$$\dot{V}_{e_{w_{i,1}}} = -\psi_{i,1}\psi_{i,1}^T w_{i,1}^T Tr(\beta_{ai,1}\hat{\omega}_{ai,1} - \frac{1}{2}\beta_{ai,1}\hat{\omega}_{ci,1} - \beta_{ci,1}\hat{\omega}_{ci,1}). \tag{4.22}$$

Taking $\beta_{i,1} = \min\{\beta_{ai,1}, \frac{1}{2}\beta_{ai,1} + \beta_{ci,1}\} \geq 0$, and letting $\lambda_{\min}(\psi_{i,1}\psi_{i,1}^T)$ denote the minimum eigenvalue of $\psi_{i,1}\psi_{i,1}^T$, we obtain:

$$\dot{V}_{e_{w_{i,1}}} \leq -\beta_{i,1}\lambda_{min}(\psi_{i,1}\psi_{i,1}^T)Tr(w_{i,1}^T w_{i,1}). \tag{4.23}$$

According to the Lyapunov stability theory, the Lyapunov function is convergent, so $\hat{\omega}_{ai,1} \to \hat{\omega}_{ci,1}$. Therefore, the selected weight update law for the critic network is appropriate.

Finally, it is also necessary to design the weight vector update law $\dot{\hat{\omega}}_{h_i,1}$ for the unknown nonlinear function $\hat{h}_{i,1}(x_{i,1})$, where both $\hat{h}_{i,1}(x_{i,1})$ and $\hat{\omega}_{h_i,1}$ are obtained after evaluation and adjustment by the critic–actor network. In this cases, $\dot{\hat{\omega}}_{h_i,1}$ is designed as

$$\dot{\hat{\omega}}_{h_i,k_1l} = \psi_{h_i,1}g_{i,1l} - \kappa_{h_{11},k_1l}\hat{\omega}_{h_i,k_1l} - \kappa_{h_{12},k_1l}\hat{\omega}_{h_i,k_1l}^3, \tag{4.24}$$

where $\kappa_{h_{11},k_1l} > 0$, $\kappa_{h_{12},k_1l} > 0$, and $\hat{\omega}_{h_i,k_1l}$ denotes the element in the k_1-th row and l-th column of $\hat{\omega}_{h_i,1}$, with $k_1 = 1, \cdots, m_{11}$ and $l = 1, \cdots, n$.

Remark 4.3. *The critic–actor network framework has been applied in works [7, 8, 25, 30], but there is no introduction to its network structure dimension or the number of neurons. Here, we mainly demonstrate it through the weight vector $\hat{\omega}_{h_i,1}$ of the unknown nonlinear function $h_{i,1}(x_{i,1})$. The advantage of this approach is that after analyzing the structural dimension of the critic–actor network, for the weight vector $\hat{\omega}_{h_i,1}$, each element composing the vector can be obtained for scalar calculation; i.e., $\hat{\omega}_{h_i,k_1l}$ is a scalar value, which only requires basic numerical operations in the calculation process. However, it should be noted that the structural dimensions of the critic–actor network are different at each backstepping step.*

Step (j, $j = 2, \cdots, n-1$). Similar to the first step, we first select the objective function $J(g_{i,j}, \dot{g}_{i,j})$, and minimizing it yields

$$J^x(g_{i,j}, \dot{g}_{i,j}) = \min_{\delta_{i,j}\in\Omega_x} \int_t^{+\infty} (g_{i,j}^2(\chi) + \dot{g}_{i,j}^2)d\chi. \tag{4.25}$$

Construct HJB equation as follows:

$$H^x(g_{i,j}, \dot{g}_{i,j}) = g_{i,j}^2 + \dot{g}_{i,j}^2 + \frac{\partial J^x(g_{i,j}, \dot{g}_{i,j})}{\partial g_{i,j}}\dot{g}_{i,j}. \tag{4.26}$$

The derived optimal control input is $\delta^x_{i,j}$=$-\frac{1}{2}\frac{\partial J^x(g_{i,j},\dot{g}_{i,j})}{\partial g_{i,j}} - \hat{\omega}^T_{i,j}\psi_{i,j} + \hat{\delta}^x_{i,j}$.The derived optimal control input is:

$$\begin{aligned}\frac{\partial J^x(g_{i,j},\dot{g}_{i,j})}{\partial g_{i,j}} =&2\eta_{i,j}(g^T_{i,j}g_{i,j})g_{i,j} + 2\nu_{i,j}(g^T_{i,j}g_{i,j})^{p-1}g_{i,j}\\ &+ 4g_{i,j} + 2h_{i,j}(x_{i,j}) + J^{x_j},\end{aligned} \tag{4.27}$$

$$\begin{aligned}J^{x_j} =& -2\eta_{i,j}(g^T_{i,j}g_{i,j})g_{i,j} - 2\nu_{i,j}(g^T_{i,j}g_{i,j})^{p-1}g_{i,j}\\ &- 4g_{i,j} - 2h_{i,j}(x_{i,j}) + \frac{\partial J^x(g_{i,j},\dot{g}_{i,j})}{\partial g_{i,j}}.\end{aligned} \tag{4.28}$$

Substituting Eqs (4.27) and (4.28) into $\delta^x_{i,j}$ yields:

$$\begin{aligned}\delta^x_{i,j} =& -\eta_{i,j}(g^T_{i,j}g_{i,j})g_{i,j} - \nu_{i,j}(g^T_{i,j}g_{i,j})^{p-1}g_{i,j} - 2g_{i,j}\\ &- h_{i,j}(x_{i,j}) - \frac{1}{2}J^{x_x} - \hat{\omega}^T_{fi,j}\psi_{fi,j} + \dot{\delta}^x_{i,j-1}.\end{aligned} \tag{4.29}$$

By approximating $h_{i,j}(x_{i,j})$ and J^{x_j}, we obtain:

$$h_{i,j}(x_{i,j}) = \omega^{*T}_{h_i,j}\psi_{h_i,j} + \epsilon_{h_i,j}, \tag{4.30}$$

$$J^{x_j} = \omega^{*T}_{J_j}\psi_{J_j} + \epsilon_{J_j}, \tag{4.31}$$

where $\omega_{h_i,j} \in \mathbb{R}^{m_{j1}\times n}$, $\psi_{h_i,j} = [\psi_{h_i,j1}, \cdots, \psi_{h_i,jm_{j1}}]^T$, $\omega^{*T}_{J_j} \in \mathbb{R}^{m_{j2}\times n}$, and $\psi_{J_j} = [\psi_{J_j,1}, \cdots, \psi_{J_j,m_{j2}}]^T$. Both $\epsilon_{h_i,j}$ and ϵ_{J_j} are bounded; i.e., $\bar{\epsilon}_{h_i,j} > 0$ and $\bar{\epsilon}_{J_j} > 0$ exist such that $||\epsilon_{h_i,j}||^2 \le \bar{\epsilon}^2_{h_i,j}$ and $||\epsilon_{J_j}||^2 \le \bar{\epsilon}^2_{J_j}$.

Next, we use the critic–actor network to evaluate and adjust the derived optimal control input $\hat{\delta}^x_{i,j}$. Similar to the first step, after evaluation and adjustment, $h_{i,j}(x_{i,j})$ and J^{x_j} are adjusted to $\hat{h}_{i,j}(x_{i,j})$ and $\hat{J}^{x_j}$. Meanwhile, the weight of the critic network $\hat{\omega}_{ci,j}$ and the weight of the actor network $\hat{\omega}_{ai,j}$ can be obtained, and their weight update laws satisfy

$$\hat{\omega}_{ai,j} = -\beta_{ai,j}\psi_{i,j}\psi^T_{i,j}(\hat{\omega}_{ai,j} - \frac{1}{2}\hat{\omega}_{ci,j}), \tag{4.32}$$

$$\dot{\hat{\omega}}_{ci,j} = -\beta_{ci,j}\psi_{i,j}\psi^T_{i,j}\hat{\omega}_{ci,j}, \tag{4.33}$$

where $\beta_{ai,j} > 0, \beta_{ci,j} > 0$, $\omega_{ai,j} \in \mathbb{R}^{m_{j2}\times n}$, and $\omega_{ci,j} \in \mathbb{R}^{m_{j2}\times n}$.

The update weight $\dot{\hat{\omega}}_{ai,j}$ of the actor network is derived using the Bellman residual, and the process is as follows:

$$E_{i,j} = H^x(g_{i,j}, \dot{\hat{g}}_{i,j}) - H^x(g_{i,j}, \dot{g}_{i,j}) = H^x(g_{i,j}, \dot{\hat{g}}_{i,j}) = 0. \tag{4.34}$$

Substituting $\frac{\partial \hat{J}^x(z_{i,1},\dot{g}_{i,1})}{\partial z_{i,1}}$ and $\hat{\delta}^x_{i,1}$ into $H^x(z_{i,1}, \dot{g}_{i,1})$, the residual is obtained as follows:

$$\begin{aligned}E_{i,j} =&g^2_{i,j} + (-\eta_{i,j}(g^T_{i,j}g_{i,j})g_{i,j} - \nu_{i,j}(g^T_{i,j}g_{i,j})^{p-1}g_{i,j}\\ &- \frac{1}{2}g_{i,j} - \omega^{*T}_{h_i,j}\psi_{h_i,j}(x_{i,j}) - \epsilon_{h_i,j} - \frac{1}{2}(\hat{\omega}^T_{ai,j}\psi_{i,j}))^2\\ &+ [2\eta_{i,j}(g^T_{i,j}g_{i,j})g_{i,j} + \frac{2\nu_{i,j}}{\pi_{i,j}}(g^T_{i,j}g_{i,j})^{p-1}g_{i,j}\\ &+ g_{i,j} + 2\omega^{*T}_{h_i,j}\psi_{h_i,j}(x_{i,j}) + 2\epsilon_{h_i,j} + (\hat{\omega}^T_{ci,j}\psi_{i,j})]\\ &[-\eta_{i,j}(g^T_{i,j}g_{i,j})g_{i,j} - \nu_{i,j}(g^T_{i,j}g_{i,j})^{p-1}g_{i,j}\\ &- \frac{1}{2}g_{i,j} - \omega^{*T}_{h_i,j}\psi_{h_i,j}(x_{i,j}) - \epsilon_{h_i,j} - \frac{1}{2}(\hat{\omega}^T_{ai,j}\psi_{i,j}) + \hat{\omega}^T_{fi,j}\psi_{fi,j}].\end{aligned} \tag{4.35}$$

Taking the partial derivative of $E_{i,j}$ with respect to $\omega_{ai,j}$, we obtain

$$\frac{\partial E_{i,j}}{\partial \omega_{ai,j}} = \frac{1}{2}\psi_{i,j}\psi_{i,j}^T(\hat{\omega}_{ai,j} - \hat{\omega}_{ci,j}) = 0. \tag{4.36}$$

Similarly, taking $W_{i,j} = \hat{\omega}_{ai,j}\frac{\partial E_{i,j}}{\partial \omega_{ai,j}}$ and setting $\frac{\partial W_{i,j}}{\partial \hat{\omega}_{ai,j}} = 0$, we obtain

$$\dot{\hat{\omega}}_{ai,j} = -\beta_{ai,j}\psi_{i,j}\psi_{i,j}^T(\hat{\omega}_{ai,j} - \frac{1}{2}\hat{\omega}_{ci,j}). \tag{4.37}$$

Select the weight vector update law of the critic network as follows:

$$\dot{\hat{\omega}}_{ci,j} = -\beta_{ci,j}\psi_{i,j}\psi_{i,j}^T\hat{\omega}_{ci,j}. \tag{4.38}$$

To prove that the critic network weight $\hat{\omega}_{ai,j}$ and the actor network weight $\hat{\omega}_{ci,j}$ tend to be consistent, we construct a Lyapunov function $V_{e_{w_{i,j}}} = \frac{1}{2}Tr(w_{i,j}^T w_{i,j})$, where $w_{i,j} = \hat{\omega}_{ai,j} - \hat{\omega}_{ci,j}$.

$$\dot{V}_{e_{w_{i,j}}} \leq -\beta_{i,j}\lambda_{min}(\psi_{i,j}\psi_{i,j}^T)Tr(w_{i,j}^T w_{i,j}). \tag{4.39}$$

Taking $\beta_{i,j} = \min\{\beta_{ai,j}, \frac{1}{2}\beta_{ai,j} + \beta_{ci,j}\} \geq 0$ and let $\lambda_{\min}(\psi_{i,j}\psi_{i,j}^T)$ denote the minimum eigenvalue of $\psi_{i,j}\psi_{i,j}^T$. According to Lyapunov stability theory, it follows that $\omega_{ai,j} \to \omega_{ci,j}$.

Finally, we design the weight vector update law $\dot{\hat{\omega}}_{h_i,j}$ for the unknown nonlinear function $\hat{h}_{i,j}(x_{i,j})$. Here, both $\hat{h}_{i,j}(x_{i,j})$ and $\hat{\omega}_{h_i,j}$ are obtained after evaluation and adjustment by the critic–actor network; i.e., obtained by evaluating and adjusting $h_{i,1}(x_{i,1})$ and $\omega^*_{h_i,j}$. The design of $\dot{\hat{\omega}}_{h_i,j}$ is as follows:

$$\dot{\hat{\omega}}_{h_i,k_jl} = \psi_{h_i,j}g_{i,jl} - \kappa_{h_{j1},k_jl}\hat{\omega}_{h_i,k_jl} - \kappa_{h_{j2},k_jl}\hat{\omega}^3_{h_i,k_jl}, \tag{4.40}$$

where $\kappa_{h_{j1},k_jl} > 0$, $\kappa_{h_{j2},k_jl} > 0$, and $\hat{\omega}_{h_j,k_jl}$ denotes the element in the k_j-th row and l-th column of $\hat{\omega}_{h_j}$, with $k_j = 1, \cdots, m_{j1}$, and $l = 1, \cdots, n$.

Step (n). The control input derived in the n-th backstepping step is denoted by u_i^x. After evaluation and adjustment by the critic–actor network, the optimal control input is: $\hat{u}_i^x$. In the n-th step, we select the objective function and minimize it, yielding

$$J^x(g_{i,n}, \dot{g}_{i,n}) = \min_{\delta_{i,j}\in\Omega_x} \int_t^{+\infty} \left(g_{i,n}^2(\chi) + \dot{g}_{i,n}^2\right) d\chi,$$

the HJB equation as follows:

$$H^x(g_{i,n}, \dot{g}_{i,n}) = g_{i,n}^2 + \dot{g}_{i,n}^2 + \frac{\partial J^x(g_{i,n}, \dot{g}_{i,n})}{\partial g_{i,n}}\dot{g}_{i,n}. \tag{4.41}$$

From $\frac{\partial H^x(g_{i,n}, \dot{g}_{i,n})}{\partial \hat{u}_i^x} = 0$, we obtain

$$2(\hat{u}_i^x + \hat{\omega}_{i,n}^T\psi_{i,n} - \dot{\delta}_{i,n-1}^x) + \frac{\partial J^x(g_{i,n}, \dot{g}_{i,n})}{\partial g_{i,n}} = 0, \tag{4.42}$$

$$u_i^x = -\frac{1}{2}\frac{\partial J^x(g_{i,n}, \dot{g}_{i,n})}{\partial g_{i,n}} - \hat{\omega}_{i,n}^T\psi_{i,n} + \dot{\delta}_{i,n-1}^x. \tag{4.43}$$

Rewrite $\frac{\partial J^x(g_{i,n},\dot{g}_{i,n})}{\partial g_{i,n}}$ as

$$\begin{aligned}\frac{\partial J^x(g_{i,n},\dot{g}_{i,n})}{\partial g_{i,n}} = & 2\eta_{i,n}(g_{i,n}^T g_{i,n})g_{i,n} + 2\nu_{i,n}(g_{i,n}^T g_{i,n})^{p-1} g_{i,n} \\ & + 4g_{i,n} + 2h_{i,n}(x_{i,n}) + J^{x_n},\end{aligned} \tag{4.44}$$

$$\begin{aligned}J^{x_n} = & - 2\eta_{i,n}(g_{i,n}^T g_{i,n})g_{i,n} - 2\nu_{i,n}(g_{i,n}^T g_{i,n})^{p-1} g_{i,n} \\ & - 4g_{i,n} - 2h_{i,n}(x_{i,n}) + \frac{\partial J^x(g_{i,n},\dot{g}_{i,n})}{\partial g_{i,n}}.\end{aligned} \tag{4.45}$$

Substituting Eq (4.44) into Eq (4.43) yields

$$\begin{aligned}u_i^x = & - \eta_{i,n}(g_{i,n}^T g_{i,n})g_{i,n} - \nu_{i,n}(g_{i,n}^T g_{i,n})^{p-1} g_{i,n} \\ & - 2g_{i,n} - h_{i,n}(x_{i,n}) - \frac{1}{2}J^{x_n} - \hat{\omega}_{i,n}^T\psi_{i,n} + \dot{\delta}_{i,n-1}^x.\end{aligned} \tag{4.46}$$

Using NNs to approximate the unknown nonlinear functions $h_{i,n}(x_{i,n})$ and J^{x_n}, we obtain

$$h_{i,n}(x_{i,n}) = \omega_{h_i,n}^{*T}\psi_{h_i,n} + \epsilon_{h_i,n}, \tag{4.47}$$

$$J^{x_n} = \omega_{J_n}^{*T}\psi_{J_n} + \epsilon_{J_n}, \tag{4.48}$$

where $\omega_{h_i,n} \in \mathbb{R}^{m_{n1}\times n}$, $\psi_{h_i,n} = [\psi_{h_i,n1}, \cdots, \psi_{h_i,nm_{n1}}]^T$, $\omega_{J_n}^{*T} \in \mathbb{R}^{m_{n2}\times n}$, and $\psi_{J_n} = [\psi_{J_n,1}, \cdots, \psi_{J_n,m_{n2}}]^T$.

Using the critic–actor network to evaluate and adjust the optimal control input u_i^x, the obtained weight vectors of the critic network and actor network, $\hat{\omega}_{ai,n}$ and $\hat{\omega}_{ci,n}$, satisfy

$$\begin{aligned}\frac{\partial J^x(g_{i,n},\dot{g}_{i,n})}{\partial g_{i,n}} = & 2\eta_{i,n}(g_{i,n}^T g_{i,n})g_{i,n} + 2\nu_{i,n}(g_{i,n}^T g_{i,n})^{p-1} g_{i,n} + g_{i,n} \\ & + 2\omega_{h_i,n}^{*T}\psi_{h_i,n} + 2\epsilon_{h_i,n} + \hat{\omega}_{ai,n}^T\psi_{ai,n},\end{aligned} \tag{4.49}$$

$$\begin{aligned}u_i^x = & - \eta_{i,n}(g_{i,n}^T g_{i,n})g_{i,n} - \nu_{i,n}(g_{i,n}^T g_{i,n})^{p-1} g_{i,n} \\ & - \frac{1}{2}g_{i,n} - \omega_{h_i,n}^{*T}\psi_{h_i,n} - \epsilon_{h_i,n} - \frac{1}{2}(\hat{\omega}_{ci,n}^T\psi_{ci,n}) \\ & - \hat{\omega}_{i,n}^T\psi_{i,n} + \dot{\delta}_{i,n-1}^x.\end{aligned} \tag{4.50}$$

The weight update laws of $\hat{\omega}_{ai,n}$ and $\hat{\omega}_{ci,n}$ satisfy

$$\dot{\hat{\omega}}_{ai,n} = -\beta_{ai,n}\psi_{i,n}\psi_{i,n}^T(\hat{\omega}_{ai,n} - \frac{1}{2}\hat{\omega}_{ci,n}), \tag{4.51}$$

$$\dot{\hat{\omega}}_{ci,n} = -\beta_{ci,n}\psi_{i,n}\psi_{i,n}^T\hat{\omega}_{ci,n}. \tag{4.52}$$

The actor network's weight update law is derived using the Bellman residual, with the main process as follows:

$$\begin{aligned}E_{i,n} = & H^x(g_{i,n},\dot{\hat{g}}_{i,n}) - H^x(g_{i,n},\dot{g}_{i,n}) \\ = & H^x(g_{i,n},\dot{\hat{g}}_{i,n}) = 0.\end{aligned} \tag{4.53}$$

Similar to the first $n-1$ steps, substituting Eqs (4.50) and (4.51) into Eq (4.42) yields $E_{i,n} = H^x(g_{i,n}, \dot{\hat{g}}_{i,n})$.

$$\begin{aligned} E_{i,j} =& g_{i,n}^2 + (-\eta_{i,n}(g_{i,n}^T g_{i,n}) g_{i,n} - \nu_{i,n}(g_{i,n}^T g_{i,n})^{p-1} g_{i,n} \\ &- \frac{1}{2} g_{i,n} - \omega_{h_i,n}^{*T} \psi_{h_i,n}(x_{i,n}) - \epsilon_{h_i,n} \\ &- \frac{1}{2}(\hat{\omega}_{ai,n}^T \psi_{i,n}))^2 + [2\eta_{i,n}(g_{i,n}^T g_{i,n}) g_{i,n} \\ &+ \frac{2\nu_{i,n}}{\pi_{i,n}}(g_{i,n}^T g_{i,n})^{p-1} g_{i,n} + g_{i,n} + 2\omega_{h_i,n}^{*T} \psi_{h_i,n}(x_{i,n}) \\ &+ 2\epsilon_{h_i,n} + (\hat{\omega}_{ci,n}^T \psi_{i,n})][-\eta_{i,n}(g_{i,n}^T g_{i,n}) g_{i,j} \\ &- \nu_{i,n}(g_{i,n}^T g_{i,n})^{p-1} g_{i,n} - \frac{1}{2} g_{i,n} - \omega_{h_i,n}^{*T} \psi_{h_i,n}(x_{i,n}) \\ &- \epsilon_{h_i,n} - \frac{1}{2}(\hat{\omega}_{ai,n}^T \psi_{i,n}) + \hat{\omega}_{fi,n}^T \psi_{fi,n}]. \end{aligned} \tag{4.54}$$

Let $W_{i,n} = \hat{\omega}_{i,n} \frac{\partial E_{i,n}}{\partial \hat{\omega}_{ai,n}}$, $\dot{\hat{\omega}}_{ai,n} = -\beta_{ai,n} \frac{\partial W_{i,n}}{\hat{\omega}_{ai,n}}$, and $\dot{\hat{\omega}}_{ai,n} = \beta_{ai,n} \psi_{i,n} \psi_{i,n}^T (\hat{\omega}_{ai,n} - \frac{1}{2}\hat{\omega}_{ci,n})$.

To prove $\hat{\omega}_{ai,n} \to \hat{\omega}_{ci,n}$ using Lyapunov stability theory, we take $w_{i,n} = \hat{\omega}_{ai,n} - \hat{\omega}_{ci,n}$ and select the Lyapunov function $V_{e_{w_{i,n}}} = \frac{1}{2} Tr(w_{i,n}^T w_{i,n})$. It follows that

$$\dot{V}_{e_{w_{i,n}}} \le -\beta_{i,n} \lambda_{min}(\psi_{i,n} \psi_{i,n}^T) Tr(w_{i,n}^T w_{i,n}), \tag{4.55}$$

where $\beta_{i,n} = \min\{\beta_{ai,n}, \frac{1}{2}\beta_{ai,n} + \beta_{ci,n}\} \ge 0$, $\dot{V}_{e_{w_{i,n}}} \le 0$, and $\hat{\omega}_{ai,n} \to \hat{\omega}_{ci,n}$.

The weight vector update law $\hat{\omega}_{h_i,n}$ for the evaluated and adjusted nonlinear function $\hat{h}_{i,n}$ is designed as:

$$\dot{\hat{\omega}}_{h_i,k_n l} = \psi_{h_i,n} g_{i,nl} - \kappa_{h_{n1},k_n l} \hat{\omega}_{h_i,k_n l} - \kappa_{h_{n2},k_n l} \hat{\omega}_{h_i,k_n l}^3, \tag{4.56}$$

$\kappa_{h_{n1},k_n l} > 0$, $\kappa_{h_{n2},k_n l} > 0$, and $\hat{\omega}_{h_i,k_n l}$ denotes the element in the k_n-th row and l-th column of $\hat{\omega}_{h_i,n}$, where $k_n = 1, \cdots, m_{n_1}$ and $l = 1, \cdots, n$.

5. Proof of fixed-time consensus

In Sections 2 and 3, the backstepping method is integrated into the observer–critic–actor network framework. The approach proceeds in two main phases. First, the optimal control input is derived; subsequently, it is evaluated and refined through the critic–actor mechanism, thereby achieving further performance enhancement. In this section, we construct Lyapunov functions using the derived optimal control input $\hat{\delta}_{i,j}^x$, the critic network weights $\hat{\omega}_{ci,j}$, the actor network weights $\hat{\omega}_{ai,j}$, and $\hat{\omega}_{h_i,k_j l}$ to prove the fixed-time consensus of the multi-agent system. We still align divide the process into n steps and construct n Lyapunov functions.

Theorem 5.1. *For the multi-agent system (2.1), if the control input $\delta_{i,j}$ satisfies (4.14), (4.29), and (4.50); the critic network weight update law $\hat{\omega}_{i,j}$ satisfies (4.19), (4.32), and (4.51); the actor network weight $\hat{\omega}_{i,j}$ satisfies (4.20), (4.33), and (4.52); the critic–actor network weight coefficients satisfy $\frac{\beta_{ai,j}}{4} > 1$, $\frac{\beta_{ai,j}}{2} > 1$, and $\omega_{h_j}^* \ge 1$; and $\omega_{h_1}^*$ is the minimum values of $\omega_{h_i,k_j l}^*$ then the fixed-time consensus of the system (1) can be achieved ($i = 1, \cdots, N$, $k_j = 1, \cdots, m_j, l = 1, \cdots, n$).*

Step (1). The Lyapunov function constructed for the first backstepping step is as follows:

$$\begin{aligned} V_1 =& \frac{1}{2}\sum_{i=1}^{N} g_{i,1}^T g_{i,1} + \sum_{i=1}^{N}\sum_{k_1=1}^{m_{11}}\sum_{l=1}^{n} \tilde{\omega}_{h_i,k_1 l}^2 \\ &+ \sum_{i=1}^{N} Tr(\tilde{\omega}_{ai,1}^T \tilde{\omega}_{ai,1}) + \sum_{i=1}^{N} Tr(\tilde{\omega}_{ci,1}^T \tilde{\omega}_{ci,1}), \end{aligned} \tag{5.1}$$

where $\tilde{\omega}_{h_i,k_1 l} = \hat{\omega}_{h_i,k_1 l} - \omega_{h_i,k_1 l}^*$, $\tilde{\omega}_{ai,1} = \hat{\omega}_{ai,1} - \omega_{i,1}^*$, and $\tilde{\omega}_{ci,1} = \hat{\omega}_{ci,1} - \omega_{i,1}^*$. Substituting $\hat{\delta}_{i,1}^x$, $\dot{\hat{\omega}}_{ai,1}$, $\dot{\hat{\omega}}_{ci,1}$, and $\dot{\hat{\omega}}_{h_i,k_1 l}$ obtained in the first step of the previous section into Eq (5.1) yields

$$\begin{aligned} \dot{V}_1 =& \sum_{i=1}^{N} g_{i,1}^T [-\eta_{i,1}(g_{i,1}^T g_{i,1}) g_{i,1} - \nu_{i,1}(g_{i,1}^T g_{i,1})^{p-1} g_{i,1} \\ &- 2g_{i,1} - \tilde{\omega}_{h_i,k_1 l}^T \psi_{h_i,1} - \epsilon_{h_i,1} - \frac{1}{2}\hat{\omega}_{ai,1}^T \psi_{i,1} - \omega_{h_i,k_1 l}^{*T} \psi_{h_i,1}] \end{aligned} \tag{5.2}$$

$$\begin{aligned} &+ \sum_{i=1}^{N}\sum_{k_1=1}^{m_{11}}\sum_{l=1}^{n} \tilde{\omega}_{h_i,k_1 l}(\psi_{h_i,1} g_{i,1l} + \kappa_{h_{11},k_1 l}\hat{\omega}_{h_i,k_1 l} \\ &- \kappa_{h_{12},k_1 l}\hat{\omega}_{h_i,k_1 l}^2) + \sum_{i=1}^{N} Tr(\tilde{\omega}_{ci,1}^T \beta_{ci,1}\psi_{i,1}\psi_{i,1}^T \hat{\omega}_{ci,1}) \\ &+ \sum_{i=1}^{N} Tr[\tilde{\omega}_{ai,1}^T \beta_{ai,1}\psi_{i,1}\psi_{i,1}^T(\hat{\omega}_{ai,1} - \frac{1}{2}\hat{\omega}_{ci,1})]. \end{aligned}$$

Using Young's inequality and the method in [14], the following formula can be obtained:

$$g_{i,1}^T \epsilon_{h_i,1} \le \frac{1}{2} g_{i,1}^T g_{i,1} + \frac{1}{2}\bar{\epsilon}_{h_i,1}^2, \tag{5.3}$$

$$g_{i,1}^T \omega_{h_i,k_1 l}^{*T} \psi_{h_i,1} \le \frac{1}{2} g_{i,1}^T g_{i,1} + \frac{1}{2} Tr(\omega_{h_i,k_1 l}^{*T} \psi_{h_i,1}\psi_{h_i,1}^T \omega_{h_i,k_1 l}^*), \tag{5.4}$$

$$g_{i,1}^T \hat{\omega}_{ai,1}^T \psi_{i,1} \le \frac{1}{2} g_{i,1}^T g_{i,1} + \frac{1}{2}\hat{\omega}_{ai,1}^T \psi_{i,1}\psi_{i,1}^T \hat{\omega}_{ai,1}, \tag{5.5}$$

$$\tilde{\omega}_{ai,1}^T \psi_{i,1}\psi_{i,1}^T \hat{\omega}_{ai,1} = \frac{1}{2}\omega_{i,1}^{*T}\psi_{i,1}\psi_{i,1}^T \omega_{i,1}^* - \frac{1}{2}\tilde{\omega}_{ai,1}^T \psi_{i,1}\psi_{i,1}^T \tilde{\omega}_{ai,1} - \frac{1}{2}\hat{\omega}_{ai,1}^T \psi_{i,1}\psi_{i,1}^T \hat{\omega}_{ai,1}, \tag{5.6}$$

$$\tilde{\omega}_{ci,1}^T \psi_{i,1}\psi_{i,1}^T \hat{\omega}_{ci,1} = \frac{1}{2}\omega_{i,1}^{*T}\psi_{i,1}\psi_{i,1}^T \omega_{i,1}^* - \frac{1}{2}\tilde{\omega}_{ci,1}^T \psi_{i,1}\psi_{i,1}^T \tilde{\omega}_{ci,1} - \frac{1}{2}\hat{\omega}_{ci,1}^T \psi_{i,1}\psi_{i,1}^T \hat{\omega}_{ci,1}, \tag{5.7}$$

$$\tilde{\omega}_{ai,1}^T \psi_{i,1}\psi_{i,1}^T \hat{\omega}_{ci,1} \le \frac{1}{2}\tilde{\omega}_{ai,1}^T \psi_{i,1}\psi_{i,1}^T \tilde{\omega}_{ai,1} + \frac{1}{2}\hat{\omega}_{ci,1}^T \psi_{i,1}\psi_{i,1}^T \hat{\omega}_{ci,1}, \tag{5.8}$$

$$g_{i,1}^T \tilde{\omega}_{h_i,k_1 l}^T \psi_{h_i,1} = \sum_{k_1=1}^{m_{11}}\sum_{i=1}^{n} \tilde{\omega}_{h_i,k_1 l}\psi_{h_i,1} g_{i,1l}. \tag{5.9}$$

Substituting Eqs (5.3)–(5.9) into Eq (5.2) yields

$$
\begin{aligned}
\dot{V}_1 \leq & \sum_{i=1}^{N}[-\eta_{i,1}(g_{i,1}^T g_{i,1})^2 - \nu_{i,1}(g_{i,1}^T g_{i,1})^p] \\
& + \sum_{i=1}^{N}\sum_{k_1=1}^{m_{11}}\sum_{l=1}^{n}\tilde{\omega}_{h_i,k_1 l}(\kappa_{h_{11},k_1 l}\hat{\omega}_{h_i,k_1 l} - \kappa_{h_{12},k_1 l}\hat{\omega}_{h_i,k_1 l}^3) \\
& - \sum_{i=1}^{N}\frac{\beta_{ai,1}}{4}Tr(\tilde{\omega}_{ai,1}^T\psi_{i,1}\psi_{i,1}^T\tilde{\omega}_{ai,1}) - \sum_{i=1}^{N}\frac{\beta_{ci,1}}{2}Tr(\tilde{\omega}_{ci,1}^T\psi_{i,1}\psi_{i,1}^T\tilde{\omega}_{ci,1}) \\
& + \sum_{i=1}^{N}\frac{1}{2}\omega_{h_i,k_1 l}^{*T}\psi_{h_i,1}\psi_{h_i,1}^T\omega_{h_i,k_1 l}^{*} + \sum_{i=1}^{N}(\frac{\beta_{ai,1}+\beta_{ci,1}}{2})Tr(\omega_{i,1}^{*T}\psi_{i,1}\psi_{i,1}^T\omega_{i,1}^{*}).
\end{aligned}
\tag{5.10}
$$

Furthermore, we prove that $\kappa_{h_{11},k_1 l}\tilde{\omega}_{h_i,k_1 l} - \kappa_{h_{12},k_1 l}\tilde{\omega}_{h_i,k_1 l}\hat{\omega}_{h_i,k_1 l}^3$ can be scaled as follows:

$$
\begin{aligned}
& \kappa_{h_{11},k_1 l}\tilde{\omega}_{h_i,k_1 l}\hat{\omega}_{h_1,k_1 l} - \kappa_{h_{12},k_1 l}\tilde{\omega}_{h_i,k_1 l}\hat{\omega}_{h_1,k_1 l}^3 \\
& \leq \frac{\kappa_{h_{11},k_1 l}}{2}\omega_{h_i,k_1 l}^{*2} - \frac{\kappa_{h_{11},k_1 l}}{2}\hat{\omega}_{h_i,k_1 l}^2 - \frac{\kappa_{h_{11},k_1 l}}{2}\tilde{\omega}_{h_i,k_1 l}^2 \\
& - \kappa_{h_{12},k_1 l}(\omega_{h_i,k_1 l}^{*3}\tilde{\omega}_{h_i,k_1 l} - 3\omega_{h_i,k_1 l}^{*2}\tilde{\omega}_{h_i,k_1 l}^2) \\
& - \kappa_{h_{12},k_1 l}(3\omega_{h_i,k_1 l}^{*}\tilde{\omega}_{h_i,k_1 l}^3 - \tilde{\omega}_{h_i,k_1 l}^4).
\end{aligned}
\tag{5.11}
$$

According to Young's inequality and the condition $\frac{\kappa_{h_{11},k_1 l}}{2} - 3\omega_{h_i,k_1 l}^{*2}\kappa_{h_{12},k_1 l} - \frac{1}{2}\kappa_{h_{12},k_1 l} - 1 \geq 0$, we can obtain

$$
\begin{aligned}
& \kappa_{h_{11},k_1 l}\tilde{\omega}_{h_i,k_1 l}\hat{\omega}_{h_i,k_1 l} - \kappa_{h_{12},k_1 l}\tilde{\omega}_{h_i,k_1 l}\hat{\omega}_{h_1,k_1 l}^3 \\
& \leq \frac{\kappa_{h_{11},k_1 l}}{2}\omega_{h_i,k_1 l}^{*2} - \frac{\kappa_{h_{11},k_1 l}}{2}\hat{\omega}_{h_i,k_1 l}^2 - \frac{\kappa_{h_{11},k_1 l}}{2}\tilde{\omega}_{h_i,k_1 l}^2 \\
& - \kappa_{h_{12},k_1 l}\omega_{h_i,k_1 l}^{*3} - 3\kappa_{h_{12},k_1 l}\omega_{h_i,k_1 l}^{*2}\tilde{\omega}_{h_i,k_1 l} \\
& + 3\kappa_{h_{12},k_1 l}\omega_{h_i,k_1 l}^{*}\tilde{\omega}_{h_i,k_1 l}^2 - \kappa_{h_{12},k_1 l}\tilde{\omega}_{h_i,k_1 l}^3 \\
& \leq -(3\kappa_{h_{12},k_1 l}\omega_{h_i,k_1 l}^{*} - \kappa_{h_{12},k_1 l})\tilde{\omega}_{h_i,k_1 l}^4 + \frac{\kappa_{h_i,k_1 l}}{2}\omega_{h_i,k_1 l}^{*6} \\
& + \kappa_{h_{12},k_1 l}\omega_{h_i,k_1 l}^{*3} + \frac{\kappa_{h_{11},k_1 l}}{2}\omega_{h_i,k_1 l}^{*2}.
\end{aligned}
\tag{5.12}
$$

Furthermore, it is necessary to prove that $\hat{\omega}_{ai,1}$ and $\hat{\omega}_{ci,1}$ are bounded. We select $V_{i,1} = \frac{1}{2}[Tr(\hat{\omega}_{ai,1}^T\hat{\omega}_{ai,1}) + Tr(\hat{\omega}_{ci,1}^T\hat{\omega}_{ci,1})]$

$$
\begin{aligned}
\dot{V}_{i,1} \leq & -\frac{3}{4}\beta_{ai,1}\lambda_{min}(\psi_{i,1}\psi_{i,1}^T)Tr(\hat{\omega}_{ai,1}^T\hat{\omega}_{ai,1}) \\
& - (\beta_{ci,1} - \frac{\beta_{ai,1}}{4})\lambda_{min}(\psi_{i,1}\psi_{i,1}^T)Tr(\hat{\omega}_{ci,1}^T\hat{\omega}_{ci,1}),
\end{aligned}
\tag{5.13}
$$

let β_i be the minimum of $\frac{3}{2}\beta_{ai,1}\lambda_{\min}(\psi_{i,1}\psi_{i,1}^T)$ and $2\beta_{ci,1}\frac{\beta_{ai,1}}{2}\lambda_{\min}(\psi_{i,1}\psi_{i,1}^T)$ Then $\dot{V}_{i,1} \leq -\beta_i V_{i,1}$. Here, $\omega_{ai,1}$ and $\omega_{ci,1}$ are decreasing and tend to converge; that is, both are bounded and smaller than their initial values. There is a $p_{i,1}(0) \geq 0$ such that $Tr(\hat{\omega}_{ai,1}^T\hat{\omega}_{ai,1}) \leq p_{i,1}(0)$ and $Tr(\hat{\omega}_{ci,1}^T\hat{\omega}_{ci,1}) \leq p_{i,1}(0)$ hold. Since $Tr(\tilde{\omega}_{si,1}^T\tilde{\omega}_{si,1}) \leq 2Tr(\omega_{si,1}^{*T}\omega_{si,1}^{*}) + Tr(\hat{\omega}_{si,1}^T\hat{\omega}_{si,1})$ $(s = a, c)$, the following inequality holds:

$$
[Tr(\tilde{\omega}_{si,1}^T\psi_{i,1}\psi_{i,1}^T\tilde{\omega}_{si,1})]^2 \leq G_{i,1},
\tag{5.14}
$$

$$G_{i,1} = 4\lambda_{max}[(\psi_{i,1}\psi_{i,1}^T)^2][Tr^2(\omega_{i,1}^{*T}\omega_{i,1}^*) + p_{i,1}^2(0) + 2Tr(\omega_{i,1}^{*T}\omega_{i,1}^*)p_{i,1}(0)]. \tag{5.15}$$

From Lemma 2.1, it is known that $p \in (0,1)$ exists such that the following inequality holds:

$$^p \leq \lambda_{max}[(\psi_{i,1}\psi_{i,1}^T)^p]Tr(\tilde{\omega}_{si,1}^T\tilde{\omega}_{si,1}) + (1-p)p^{\frac{1-p}{p}}, \tag{5.16}$$

$$\tilde{\omega}_{h_i,k_1l}^{2p} \leq \tilde{\omega}_{h_i,k_1l}^2 + (1-p)p^{\frac{1-p}{p}}. \tag{5.17}$$

In summary, on the basis of the results from Eqs (5.12) to (5.17), we can obtain

$$\begin{aligned}\dot{V}_1 \leq &-\sum_{i=1}^N \eta_{i,1}(g_{i,1}^T g_{i,1})^2 - \sum_{i=1}^N \nu_{i,1}(g_{i,1}^T g_{i,1})^p - \sum_{i=1}^N\sum_{k_1=1}^{m_{11}}\sum_{l=1}^n (3\kappa_{h_{12},k_1l}\omega_{h_i,k_1l}^* - \kappa_{h_{12},k_1l})\tilde{\omega}_{h_i,k_1l}^4 \\ &-\sum_{i=1}^N\sum_{k_1=1}^{m_{11}}\sum_{l=1}^n \tilde{\omega}_{h_i,k_1l}^{2p} - \sum_{i=1}^N \lambda_{min}[(\psi_{i,1}\psi_{i,1}^T)^p][Tr(\tilde{\omega}_{ai,1}^T\tilde{\omega}_{ai,1})]^p - \sum_{i=1}^N \lambda_{min}[(\psi_{i,1}\psi_{i,1}^T)^p][Tr(\tilde{\omega}_{ci,1}^T\tilde{\omega}_{ci,1})]^p \\ &-\sum_{i=1}^N(\frac{\beta_{ai,1}}{4} - 1)\lambda_{min}(\psi_{i,1}\psi_{i,1}^T)Tr(\tilde{\omega}_{ai,1}^T\tilde{\omega}_{ai,1}) - \sum_{i=1}^N(\frac{\beta_{ci,1}}{2} - 1)\lambda_{min}(\psi_{i,1}\psi_{i,1}^T)Tr(\tilde{\omega}_{ci,1}^T\tilde{\omega}_{ci,1}) \\ &-\sum_{i=1}^N \lambda_{main}[(\psi_{i,1}\psi_{i,1}^T)^2][Tr\tilde{\omega}_{ai,1}^T\tilde{\omega}_{ai,1})]^2 - \sum_{i=1}^N \lambda_{main}[(\psi_{i,1}\psi_{i,1}^T)^2][Tr(\tilde{\omega}_{ci,1}^T\tilde{\omega}_{ci,1})]^2 + \Gamma_1.\end{aligned} \tag{5.18}$$

Herein, in the process above, we assume that $\frac{\kappa_{h_i,k_1l}}{2}\omega_{h_i,k_1l}^{*6} + \kappa_{h_{12},k_1l}\omega_{h_i,k_1l}^{*3} + \frac{\kappa_{h_{11},k_1l}}{2}\omega_{h_i,k_1l}^{*2} = 2G_{i,1}$.

On the left-hand side of the inequality, we simultaneously add and subtract $\sum_{i=1}^n\left(\tilde{\omega}_{si,1}^T\psi_{i,1}\psi_{i,1}^T\tilde{\omega}_{si,1}\right)^p$ (s=a,c) and $\sum_{i=1}^N\sum_{k_1=1}^{m_{11}}\sum_{l=1}^n\tilde{\omega}_{h_i,k_1l}^{2p}$, then perform scaling on the basis of Eqs (5.16) and (5.17). $\Gamma_1 = \sum_{i=1}^N \frac{\beta_{ai,1}+\beta_{ci,1}}{2}\lambda_{\max}(\psi_{i,1}\psi_{i,1}^T)Tr(\omega_{i,1}^{*T}\omega_{i,1}^*) + \frac{1}{2}\sum_{i=1}^N\bar{\epsilon}_{hi,1} + 2(1-p)p^{\frac{1-p}{p}} + 2G_{i,1}$, so we organize Eq (5.18) according to Lemma 2.2 to obtain

$$\dot{V}_1 \leq -A_1V_1^2 - B_1V_1^p + \Gamma_1, \tag{5.19}$$

where $A_1 = min\{4\eta_1, \frac{4}{nm_{11}N}[3\kappa_{h_{12}}(\omega_{h_1}^* - 1)], 4N^{-1}\lambda_{min}(\psi_1\psi_1^2\}$ $B_1 = min\{2^p\nu_1, 2^p, 2^p\lambda_{min}(\psi_1\psi_1^T)\}$,$\eta_1$,$\kappa_{h_{12}}$,$\omega_{h_1}^*$ and $\lambda_{min}(\psi_1\psi_1^2)$ are the minimum values of $4\eta_1$,κ_{h_{12},k_1l},ω_{h_i,k_1l}^* and $\lambda_{main}(\psi_{i,1}\psi_{i,1}^T)^2$ $(i = 1,\cdots,N, k_1 = 1,\cdots,m_{11}, l = 1,\cdots,n)$. Here, ν_1 is the minimum of $\nu_{i,1}(i = 1,\cdots,N)$.

The proof of the reverse derivation process for the first step is completed.

Step ($j, j = 2,\cdots,n-1$)**.** The Lyapunov function is constructed as follows:

$$\begin{aligned}V_j =&\frac{1}{2}\sum_{i=1}^N g_{i,j}^T g_{i,j} + \sum_{i=1}^N\sum_{k_j=1}^{m_{j1}}\sum_{l=1}^n \tilde{\omega}_{h_i,k_jl}^2 \\ &+\sum_{i=1}^N Tr(\tilde{\omega}_{ai,j}^T\tilde{\omega}_{ai,j}) + \sum_{i=1}^N Tr(\tilde{\omega}_{ci,j}^T\tilde{\omega}_{ci,j}).\end{aligned} \tag{5.20}$$

Among them, $\tilde{\omega}_{h_i,k_jl} = \hat{\omega}_{h_i,k_jl} - \omega_{h_{i_j},k_jl}^*$, $\tilde{\omega}_{ai,j} = \hat{\omega}_{ai,j} - \omega_{i,j}^*$, and $\tilde{\omega}_{ci,j} = \hat{\omega}_{ci,j} - \omega_{i,j}^*$. Substituting $\hat{\delta}_{i,j}^x$, $\dot{\hat{\omega}}_{ai,j}$,

$\dot{\hat{\omega}}_{ci,j}$, and $\dot{\hat{\omega}}_{h_i,k_jl}$ obtained in the first step of the previous section into Eq (5.20), we can get

$$\begin{aligned}\dot{V}_j =& \sum_{i=1}^{N} g_{i,j}^T[-\eta_{i,j}(g_{i,j}^T g_{i,j})g_{i,j} - \nu_{i,j}(g_{i,j}^T g_{i,j})^{p-1} g_{i,j} \\ &- 2g_{i,j} - \tilde{\omega}_{h_i,k_jl}^T \psi_{h_i,j} - \epsilon_{h_i,j} - \frac{1}{2}\hat{\omega}_{ai,j}^T \psi_{i,j} - \omega_{h_i,k_jl}^{*T}\psi_{h_i,j}] \\ &+ \sum_{i=1}^{N}\sum_{k_j=1}^{m_{j1}}\sum_{l=1}^{n} \tilde{\omega}_{h_i,k_jl}(\psi_{h_i,j} g_{i,jl} + \kappa_{h_{j1},k_jl}\hat{\omega}_{h_i,k_1l} - \kappa_{h_{j2},k_jl}\hat{\omega}_{h_i,k_jl}^2) \\ &+ \sum_{i=1}^{N} Tr[\tilde{\omega}_{ai,j}^T \beta_{ai,j}\psi_{i,j}\psi_{i,j}^T(\hat{\omega}_{ai,j} - \frac{1}{2}\hat{\omega}_{ci,j})] + \sum_{i=1}^{N} Tr(\tilde{\omega}_{ci,j}^T \beta_{ci,j}\psi_{i,j}\psi_{i,j}^T \hat{\omega}_{ci,j}).\end{aligned} \tag{5.21}$$

According to Young's inequality, we can obtain

$$g_{i,j}^T \epsilon_{h_i,j} \leq \frac{1}{2} g_{i,j}^T g_{i,j} + \frac{1}{2}\bar{\epsilon}_{h_i,j}^2, \tag{5.22}$$

$$g_{i,j}^T \omega_{h_i,k_jl}^{*T}\psi_{h_i,j} \leq \frac{1}{2} g_{i,j}^T g_{i,j} + \frac{1}{2} Tr(\omega_{h_i,k_jl}^{*T}\psi_{h_i,j}\psi_{h_i,j}^T \omega_{h_i,k_jl}^{*}), \tag{5.23}$$

$$g_{i,j}^T \hat{\omega}_{ai,j}^T \psi_{i,j} \leq \frac{1}{2} g_{i,j}^T g_{i,j} + \frac{1}{2}\hat{\omega}_{ai,j}^T \psi_{i,j}\psi_{i,j}^T \hat{\omega}_{ai,j}, \tag{5.24}$$

$$\tilde{\omega}_{ai,j}^T \psi_{i,j}\psi_{i,j}^T \hat{\omega}_{ai,j} = \frac{1}{2}\omega_{i,j}^{*T}\psi_{i,j}\psi_{i,j}^T \omega_{i,j}^{*} - \frac{1}{2}\tilde{\omega}_{ai,j}^T \psi_{i,j}\psi_{i,j}^T \tilde{\omega}_{ai,j} - \frac{1}{2}\hat{\omega}_{ai,j}^T \psi_{i,j}\psi_{i,j}^T \hat{\omega}_{ai,j}, \tag{5.25}$$

$$\tilde{\omega}_{ci,j}^T \psi_{i,j}\psi_{i,j}^T \hat{\omega}_{ci,j} = \frac{1}{2}\omega_{i,j}^{*T}\psi_{i,j}\psi_{i,j}^T \omega_{i,j}^{*} - \frac{1}{2}\tilde{\omega}_{ci,j}^T \psi_{i,j}\psi_{i,j}^T \tilde{\omega}_{ci,j} - \frac{1}{2}\hat{\omega}_{ci,j}^T \psi_{i,j}\psi_{i,j}^T \hat{\omega}_{ci,j}, \tag{5.26}$$

$$\tilde{\omega}_{ai,j}^T \psi_{i,j}\psi_{i,j}^T \hat{\omega}_{ci,j} \leq \frac{1}{2}\tilde{\omega}_{ai,j}^T \psi_{i,j}\psi_{i,j}^T \tilde{\omega}_{ai,j} + \frac{1}{2}\hat{\omega}_{ci,j}^T \psi_{i,j}\psi_{i,j}^T \hat{\omega}_{ci,j}, \tag{5.27}$$

$$g_{i,j}^T \tilde{\omega}_{h_i,k_jl}^T \psi_{h_i,j} = \sum_{k_j=1}^{m_{j1}}\sum_{i=1}^{n} \tilde{\omega}_{h_i,k_jl}\psi_{h_i,j} g_{i,jl}. \tag{5.28}$$

In Step 1, we have already derived the scaled form of $\kappa_{h_{11},k_1l}\tilde{\omega}_{h_i,k_1l}\hat{\omega}_{h_1,k_1l} - \kappa_{h_{12},k_1l}\tilde{\omega}_{h_i,k_1l}\hat{\omega}_{h_i,k_1l}^3$. According to the results in the first step, we can obtain

$$\begin{aligned}&\kappa_{h_{j1},k_jl}\tilde{\omega}_{h_i,k_jl}\hat{\omega}_{h_i,k_jl} - \kappa_{h_{j2},k_jl}\tilde{\omega}_{h_i,k_jl}\hat{\omega}_{h_i,k_jl}^3 \\ &\leq -(3\kappa_{h_{j2},k_jl}\omega_{h_i,k_jl}^{*} - \kappa_{h_{j2},k_jl})\tilde{\omega}_{h_i,k_jl}^4 + \frac{\kappa_{h_i,k_jl}}{2}\omega_{h_i,k_jl}^{*6} \\ &+ \kappa_{h_{j2},k_jl}\omega_{h_i,k_jl}^{*3} + \frac{\kappa_{h_{j1},k_jl}}{2}\omega_{h_i,k_jl}^{*2}.\end{aligned} \tag{5.29}$$

Furthermore, according to the conclusion from the reverse derivation process of the first step, we can conclude that both $Tr(\hat{\omega}_{ai,j}^T \hat{\omega}_{ai,j})$ and $Tr(\hat{\omega}_{ci,j}^T \hat{\omega}_{ci,j})$ are bounded, i.e., $p_{i,j}(0) > 0$ exists such that $Tr(\hat{\omega}_{ai,j}^T \hat{\omega}_{ai,j}) \leq p_{i,j}(0)$, $Tr(\hat{\omega}_{ai,j}^T \hat{\omega}_{ai,j}) \leq p_{i,j}(0)$, and the following Young's inequality holds:

$$[Tr(\tilde{\omega}_{si,j}^T \psi_{i,j}\psi_{i,j}^T \tilde{\omega}_{si,j})]^2 \leq G_{i,j}, \tag{5.30}$$

where $G_{i,j} = 4\lambda_{max}[(\psi_{i,j}\psi_{i,j}^T)^2][Tr^2(\omega_{i,j}^{*T}\omega_{i,j}^*) + p_{i,j}(0)^2 + 2Tr(\omega_{i,j}^{*T}\omega_{i,j}^*)p_{i,j}(0)]$.

It follows from Lemma 2.1 that $p \in (0,1)$ exist such that the following inequality holds:

$$[Tr(\tilde{\omega}_{si,j}^T\psi_{i,j}\psi_{i,j}^T\tilde{\omega}_{si,j})]^p \leq \lambda_{max}[(\psi_{i,j}\psi_{i,j}^T)^p]Tr(\tilde{\omega}_{si,j}^T\tilde{\omega}_{si,j}) + (1-p)p^{\frac{1-p}{p}}, \tag{5.31}$$

$$\tilde{\omega}_{h_i,k_jl}^{2p} \leq \tilde{\omega}_{h_i,k_jl}^2 + (1-p)p^{\frac{1-p}{p}}. \tag{5.32}$$

Substituting the results above into Eq (5.21), we can obtain

$$\begin{aligned}\dot{V}_j \leq &-\sum_{i=1}^N \eta_{i,j}(g_{i,j}^Tg_{i,j})^2 - \sum_{i=1}^N \nu_{i,j}(g_{i,j}^Tg_{i,j})^p - \sum_{i=1}^N\sum_{k_j=1}^{m_{j1}}\sum_{l=1}^n(3\kappa_{h_{j2},k_jl}\omega_{h_i,k_jl}^* - \kappa_{h_{j2},k_jl})\tilde{\omega}_{h_i,k_jl}^4 \\ &-\sum_{i=1}^N\sum_{k_j=1}^{m_{j1}}\sum_{l=1}^n \tilde{\omega}_{h_i,k_jl}^{2p} - \sum_{i=1}^N \lambda_{min}(\psi_{i,j}\psi_{i,j}^T)^p[Tr(\tilde{\omega}_{ai,j}^T\tilde{\omega}_{ai,j})]^p - \sum_{i=1}^N \lambda_{min}(\psi_{i,j}\psi_{i,j}^T)^P[Tr(\tilde{\omega}_{ci,j}^T\tilde{\omega}_{ci,j})]^p \\ &-\sum_{i=1}^N(\frac{\beta_{ai,j}}{4}-1)\lambda_{min}(\psi_{i,j}\psi_{i,j}^T)Tr(\tilde{\omega}_{ai,j}^T\tilde{\omega}_{ai,j}) - \sum_{i=1}^N(\frac{\beta_{ci,j}}{2}-1)\lambda_{min}(\psi_{i,j}\psi_{i,j}^T)Tr(\tilde{\omega}_{ci,j}^T\tilde{\omega}_{ci,j}) \\ &-\sum_{i=1}^N \lambda_{main}[(\psi_{i,j}\psi_{i,j}^T)^2][Tr(\tilde{\omega}_{ai,j}^T\tilde{\omega}_{ai,j})]^2 - \sum_{i=1}^N \lambda_{main}[(\psi_{i,j}\psi_{i,j}^T)^2][Tr(\tilde{\omega}_{ci,j}^T\tilde{\omega}_{ci,j})]^2 + \Gamma_j.\end{aligned} \tag{5.33}$$

By rearranging Eq (5.33), we can obtain

$$\dot{V}_j \leq -A_jV_j^2 - B_jV_j^p + \Gamma_j, \tag{5.34}$$

where $A_j = min\{4\eta_j, \frac{4}{nm_{j1}N}[3\kappa_{h_{j2}}(\omega_{h_j}^* - 1)], 4N^{-1}\lambda_{min}(\psi_j\psi_j^2)\}$ $B_j = min\{2^p\nu_j, 2^p, 2^p\lambda_{min}(\psi_j\psi_j^T)\}$, η_j, $\kappa_{h_{j2}}$, $\omega_{h_j}^*$, and$\lambda_{min}(\psi_j\psi_j^2)$ are the minimum values of $4\eta_j$, κ_{h_{j2},k_jl}, ω_{h_i,k_jl}^* and $\lambda_{main}(\psi_{i,j}\psi_{i,j}^T)^2$ $(i = 1,\cdots,N, k_j = 1,\cdots,m_{j1}, l = 1,\cdots,n)$; ν_j is the minimum of $\nu_{i,j}(i = 1,\cdots,N)$.

Step (n). The backward induction process for the n-th step is similar to that for the first $n-1$ steps. By constructing a Lyapunov function and using Young's inequality to bound its derivative, we obtain the form of Eq (2.5) in Lemma 2.3. The main process is as follows:

$$\begin{aligned}V_n =&\frac{1}{2}\sum_{i=1}^N g_{i,n}^Tg_{i,n} + \sum_{i=1}^N\sum_{k_n=1}^{m_{n1}}\sum_{l=1}^n \tilde{\omega}_{h_i,k_nl}^2 \\ &+ \sum_{i=1}^N Tr(\tilde{\omega}_{ai,n}^T\tilde{\omega}_{ai,n}) + \sum_{i=1}^N Tr(\tilde{\omega}_{ci,n}^T\tilde{\omega}_{ci,n}),\end{aligned} \tag{5.35}$$

where $\tilde{\omega}_{h_i,k_nl} = \hat{\omega}_{h_i,k_nl} - \omega_{h_i,k_nl}^*$, $\tilde{\omega}_{ai,n} = \hat{\omega}_{ai,n} - \omega_{i,n}^*$, and $\tilde{\omega}_{ci,n} = \hat{\omega}_{ci,n} - \omega_{i,n}^*$. Substituting $\hat{u}_i^x$, $\dot{\hat{\omega}}_{ai,n}$, $\dot{\hat{\omega}}_{ci,n}$, and $\dot{\hat{\omega}}_{h_i,k_nl}$ into Eq (5.35), we can obtain

$$
\begin{aligned}
\dot{V}_n =& \sum_{i=1}^{N} g_{i,n}^T[-\eta_{i,j}(g_{i,n}^T g_{i,n})g_{i,n} - \nu_{i,n}(g_{i,n}^T g_{i,n})^{p-1} g_{i,n} \\
& - 2g_{i,n} - \tilde{\omega}_{h_i,k_nl}^T \psi_{h_i,n} - \epsilon_{h_i,n} - \frac{1}{2}\hat{\omega}_{ai,n}^T \psi_{i,n} - \omega_{h_i,k_nl}^{*T}\psi_{h_i,n}] \\
& + \sum_{i=1}^{N}\sum_{k_n=1}^{m_{n1}}\sum_{l=1}^{n} \tilde{\omega}_{h_i,k_nl}(\psi_{h_i,n} g_{i,jl} + \kappa_{h_{n1},k_nl}\hat{\omega}_{h_i,k_nl} \\
& - \kappa_{h_{n2},k_nl}\hat{\omega}_{h_i,k_nl}^2) + \sum_{i=1}^{N} Tr[\tilde{\omega}_{ai,n}^T \beta_{ai,n}\psi_{i,n}\psi_{i,n}^T(\hat{\omega}_{ai,n} - \frac{1}{2}\hat{\omega}_{ci,n})] \\
& + \sum_{i=1}^{N} Tr(\tilde{\omega}_{ci,n}^T \beta_{ci,n}\psi_{i,n}\psi_{i,n}^T \hat{\omega}_{ci,n}).
\end{aligned}
\tag{5.36}
$$

According to Young's inequality, we can obtain

$$g_{i,n}^T \epsilon_{h_i,n} \leq \frac{1}{2} g_{i,n}^T g_{i,n} + \frac{1}{2}\bar{\epsilon}_{h_i,n}^2, \tag{5.37}$$

$$g_{i,n}^T \omega_{h_i,k_nl}^{*T}\psi_{h_i,n} \leq \frac{1}{2} g_{i,n}^T g_{i,n} + \frac{1}{2} Tr(\omega_{h_i,k_nl}^{*T}\psi_{h_i,n}\psi_{h_i,n}^T\omega_{h_i,k_nl}^{*}), \tag{5.38}$$

$$g_{i,n}^T \hat{\omega}_{ai,n}^T \psi_{i,n} \leq \frac{1}{2} g_{i,n}^T g_{i,n} + \frac{1}{2}\hat{\omega}_{ai,n}^T \psi_{i,n}\psi_{i,n}^T\hat{\omega}_{ai,n}, \tag{5.39}$$

$$\tilde{\omega}_{ai,n}^T \psi_{i,n}\psi_{i,n}^T \hat{\omega}_{ai,n} = \frac{1}{2}\omega_{i,n}^{*T}\psi_{i,n}\psi_{i,n}^T\omega_{i,n}^{*} - \frac{1}{2}\tilde{\omega}_{ai,n}^T\psi_{i,n}\psi_{i,n}^T\tilde{\omega}_{ai,n} - \frac{1}{2}\hat{\omega}_{ai,n}^T\psi_{i,n}\psi_{i,n}^T\hat{\omega}_{ai,n}, \tag{5.40}$$

$$\tilde{\omega}_{ci,n}^T \psi_{i,n}\psi_{i,n}^T \hat{\omega}_{ci,n} = \frac{1}{2}\omega_{i,n}^{*T}\psi_{i,n}\psi_{i,n}^T\omega_{i,n}^{*} - \frac{1}{2}\tilde{\omega}_{ci,n}^T\psi_{i,n}\psi_{i,n}^T\tilde{\omega}_{ci,n} - \frac{1}{2}\hat{\omega}_{ci,n}^T\psi_{i,n}\psi_{i,n}^T\hat{\omega}_{ci,n}, \tag{5.41}$$

$$\tilde{\omega}_{ai,n}^T \psi_{i,n}\psi_{i,n}^T \hat{\omega}_{ci,n} \leq \frac{1}{2}\tilde{\omega}_{ai,n}^T\psi_{i,n}\psi_{i,n}^T\tilde{\omega}_{ai,n} + \frac{1}{2}\hat{\omega}_{ci,n}^T\psi_{i,n}\psi_{i,n}^T\hat{\omega}_{ci,n}, \tag{5.42}$$

$$g_{i,n}^T \tilde{\omega}_{h_i,k_nl}^T \psi_{h_i,n} = \sum_{k_n=1}^{m_{n1}}\sum_{i=1}^{n} \tilde{\omega}_{h_i,k_nl}\psi_{h_i,n} g_{i,nl}, \tag{5.43}$$

$$
\begin{aligned}
& \kappa_{h_{jn},k_nl}\tilde{\omega}_{h_i,k_nl}\hat{\omega}_{h_i,k_nl} - \kappa_{h_{n2},k_nl}\tilde{\omega}_{h_i,k_nl}\hat{\omega}_{h_i,k_nl}^3 \\
& \leq -(3\kappa_{h_{n2},k_nl}\omega_{h_i,k_nl}^{*} - \kappa_{h_{n2},k_nl})\tilde{\omega}_{h_i,k_nl}^4 + \frac{\kappa_{h_i,k_nl}}{2}\omega_{h_i,k_nl}^{*6} \\
& + \kappa_{h_{n2},k_nl}\omega_{h_i,k_nl}^{*3} + \frac{\kappa_{h_{n1},k_nl}}{2}\omega_{h_i,k_nl}^{*2},
\end{aligned}
\tag{5.44}
$$

$${}^{p} \leq \lambda_{max}[(\psi_{i,n}\psi_{i,n}^T)^p] Tr(\tilde{\omega}_{si,n}^T\tilde{\omega}_{si,n}) + (1-p)p^{\frac{1-p}{p}}, \tag{5.45}$$

$$\tilde{\omega}_{h_i,k_nl}^{2p} \leq \tilde{\omega}_{h_i,k_nl}^2 + (1-p)p^{\frac{1-p}{p}}, \tag{5.46}$$

$$[Tr(\tilde{\omega}_{si,n}^T\psi_{i,n}\psi_{i,n}^T\tilde{\omega}_{si,n})]^2 \leq G_{i,n}, \tag{5.47}$$

$$\frac{\kappa_{h_i,k_nl}}{2}\omega_{h_i,k_nl}^{*6} + \kappa_{h_{n2},k_jl}\omega_{h_i,k_nl}^{*3} + \frac{\kappa_{h_{n1},k_nl}}{2}\omega_{h_i,k_nl}^{*2} = 2G_{i,n}. \tag{5.48}$$

Similar to the first $n-1$ steps, let

$$G_{i,n} = 4\lambda_{max}[(\psi_{i,n}\psi_{i,n}^T)^2][Tr^2(\omega_{i,n}^{*T}\omega_{i,n}^*) + p_{i,n}(0)^2 + 2Tr(\omega_{i,n}^{*T}\omega_{i,n}^*)p_{i,n}(0)]. \quad (5.49)$$

Substitute Eqs (5.48) and (5.49) into Eq (5.36)

$$\dot{V}_n \leq -A_nV_n^2 - B_nV_n^p + \Gamma_n. \quad (5.50)$$

Here, $A_n = min\{4\eta_n, \frac{4}{nm_{n1}N}[3\kappa_{h_{n2}}(\omega_{h_n}^* - 1)], 4N^{-1}\lambda_{min}(\psi_n\psi_n^2)\}$ $B_n = min\{2^p\nu_n, 2^p, 2^p\lambda_{min}(\psi_n\psi_n^T)\}$,$\eta_n$, $\kappa_{h_{n2}}$, $\omega_{h_n}^*$, and$\lambda_{min}(\psi_n\psi_n^2)$ are the minimum values of $4\eta_n$, κ_{h_{j2},k_nl}, ω_{h_i,k_nl}^*, and $\lambda_{main}(\psi_{i,n}\psi_{i,n}^T)^2$ $(i = 1, \cdots, N, k_n = 1, \cdots, m_{n1}, l = 1, \cdots, n)$; ν_n is the minimum of $\nu_{i,n}(i = 1, \cdots, N)$.

Up to this point, the proof for the n backstepping processes has been completed. The Lyapunov function in each backstepping process satisfies $\dot{V}_j \leq -A_jV_j^2 - B_jV_j^p + \Gamma_j$ $(j = 1, 2, \ldots, n)$. Furthermore, according to Lemma 2.2, for $V = \sum_{j=1}^n V_j$, we have.

$$\dot{V} \leq -AV^2 - BV^p + \Gamma, \quad (5.51)$$

where $p\in(0, 1)$, $A = \min\{A_1, \cdots, A_n\}$, $B = \min\{B_1, \cdots, B_n\}$, and $\Gamma = \sum_{j=1}^n \Gamma_j$. Therefore, this completes the proof of Theorem 5.1. we can also see the MASs (2.1) is practically FxT stable with the convergence time $T(x_0)$ satisfies

$$T(x_0) \leq T_{max} = \frac{1}{A\kappa(1-B)} + \frac{1}{2\kappa(1-p)}. \quad (5.52)$$

Furthermore, the convergence set of V can also be obtained as follows:

$$\left\{\lim_{t\to T(x_0)} x|V \leq min\left\{A^{-\frac{1}{c}}(\frac{\Gamma}{1-\kappa})^{\frac{1}{c}}, B^{-\frac{1}{p}}(\frac{\Gamma}{1-\kappa})^{\frac{1}{p}}\right\}\right\}. \quad (5.53)$$

Remark 5.1. *According to Eq (5.52), to shorten the convergence time, the value of B and p can be reduced while the value of A is increased. Specifically, the value of B can be decreased by reducing the values of p and* $\nu_{i,j}$*, and the value of A can be increased by increasing the value of* $\eta_{i,j}$*. Since all these values need to be obtained through the RL algorithm, the RL algorithm proposed in this paper cannot only acquire the optimal control input but also shorten the convergence time by selecting appropriate coefficient values. According to Eq (5.53), we have* $\frac{1}{2}\sum_{i=1}^N g_{i,j}^T g_{i,j} \leq V \leq \min\left\{A^{-\frac{1}{2}}\left(\frac{\Gamma}{1-\kappa}\right)^{\frac{1}{2}}, B^{-\frac{1}{p}}\left(\frac{\Gamma}{1-\kappa}\right)^{\frac{1}{p}}\right\}$. *To make* $g_{i,j}$ *converge to the desired precision, the values of A and B should be as large as possible, while the value of* Γ *should be as small as possible. According to Eqs (5.18) and (5.19), it is known that increasing the values of* $\eta_{i,j}$, κ_{h_{j2},k_jl}, *and* ω_{h_i,k_jl}^* *can increase* A_j*; increasing the value of* $\nu_{i,j}$ *can increase* B_j*; and simultaneously reducing the values of* $\beta_{ai,j}$ *and* $\beta_{ci,j}$ *while increasing the value of p can reduce the value of* Γ_j*. However, increasing or decreasing the parameters above does not imply that they can be adjusted indefinitely. This is because reducing the control error typically consumes considerable control energy. Therefore, in control design applications, a trade-off must be made between achieving better control performance and managing the control effort requred.*

Remark 5.2. *Unlike the finite-time forms* $V \leq -AV - BV^p + \Gamma$ *or* $V \leq -AV^p$*, it does not require achieving a higher-order form of V (*$V \leq -AV^q - BV^p + \Gamma$, $q > 1$*). In fixed time,* V^p *can be obtained using the method for* V^p *in finite time, such as in [7, 27]. However, obtaining the higher-order term*

V^q is the greatest challenge in this paper. We need to obtain $(g_{i,j}^T g_{i,j})^2$. The term $(g_{i,j}^T g_{i,j})g_{i,j}$ obtained when deriving the optimal control input helps solve $(g_{i,j}^T g_{i,j})^2$. However, we have no way to obtain $[Tr(\tilde{\omega}_{ai,j}^T \tilde{\omega}_{ai,j})]^2$ and $[Tr(\tilde{\omega}_{ci,j}^T \tilde{\omega}_{ci,j})]^2$. In finite time, they can completely rely on their own update laws to obtain the first-order terms $\tilde{\omega}_{ci,j}^T \tilde{\omega}_{ci,j}$ and $\tilde{\omega}_{ai,j}^T \tilde{\omega}_{ai,j}$ (see [25]). To solve this problem, we first considered boundedness. It happens that $[Tr(\tilde{\omega}_{si,j}^T \tilde{\omega}_{si,j})]^2$ (s= a, c) satisfies boundedness. Thus, we considered how to obtain a positive value that serves as the upper bound of $[Tr(\tilde{\omega}_{si,j}^T \tilde{\omega}_{si,j})]^2$ $s = (a, c)$. The selected unknown function $h_{i,j}(x_{i,j}) = \omega_{h_{i,j}}^{*T} \psi_{h_{i,j}} + \epsilon_{h_{i,j}}$ precisely solves this problem. Due to the arbitrariness of $\omega_{h_{i,j}}^*$, we let a certain form of $\omega_{h_{i,j}}^*$ be the upper bound of $[Tr(\tilde{\omega}_{si,j}^T \tilde{\omega}_{si,j})]^2$ (s=a, c). This form of $\omega_{h_{i,j}}^*$ has been given in the paper. Therefore, the term V^q is solved.

6. Numerical simulation

6.1. Images of fixed-time consistency

In this section, numerical simulations are conducted to verify the effectiveness and feasibility of the proposed scheme. we propose two single-link robotic systems for this simulation.
(1) Consider the system (2.1) as follows:

$$\begin{cases} \dot{x}_{1,1} = x_{1,2} + 0.5sin(x_{1,1})cos(x_{1,1}), \\ \dot{x}_{1,2} = u_1 + 0.5sin(x_{1,1})cos(x_{1,2}), \\ y = x_{1,1}. \end{cases} \tag{6.1}$$

In contrast to the system (2.1), we set i=1 and j=1, 2. Select the leader's output as $y_d = sin(t)$. Simultaneously select the initial values as $x_1 = [x_{1,1}(0), x_{1,2}(0)]^T = [0.75, 0.8]^T$. Select the observed value of the initial value as $\hat{x}_{1,1} = 0.7$, $\hat{x}_{1,2} = 0.7$, $g_{1,1}(0) = 0.3$, $g_{1,2}(0) = 0.3$, $\hat{\omega}_{a1}(0) = [0.75, 1.2, 1.2, 1.2]^T$, $\hat{\omega}_{a2}(0) = [1.1, 1.1, 1.1, 1.1, 1.1]^T$, $\hat{\omega}_{c1}(0)$=$[0.8, 0.8, 0.8, 0.8]^T$, $\omega_{c2}(0) = [0.9, 0.9, 0.9, 0.9]^T$, $\hat{\omega}_{a1,1}(0)$=$[\hat{\omega}_{a1}(0), \hat{\omega}_{a1}(0)]^T$ $\hat{\omega}_{a1,2}(0) = [\hat{\omega}_{a2}(0), \hat{\omega}_{a2}(0)]^T$, $\hat{\omega}_{c1,1}(0) = [\hat{\omega}_{c1}(0), \hat{\omega}_{c1}(0)]^T$, $\hat{\omega}_{c1,2}(0) = [\hat{\omega}_{c2}(0), \hat{\omega}_{c2}(0)]^T$, $\hat{\omega}_{f1,1} = [1, 1, 1, 1, 1]$, and $\hat{\omega}_{f1,2} = [1, 1, 1, 1, 1]$.

Select the matrix $Q_1 = 2I$. Then the value of P_1 is obtained from the Riccati equation as follows:

$$P_1 = \begin{bmatrix} 0.1050 & 0.1000 \\ 0.1000 & 4.1000 \end{bmatrix}.$$

For the weight update law parameters of the critic–actor network, $\beta_{a1,1}$=9, $\beta_{a1,2}$=9, $\beta_{c1,1}$=13, and $\beta_{c1,2}$ = 13; for the update law parameters of the system's unknown functions,, $\theta_{1,1}$=20, $\theta_{1,2}$=20; $\eta_{1,2} = 30$, $\nu_{1,2} = 0.1$. The observer gain coefficients are designed as $w_{1,1} = w_{1,2} = 40$.

Based on the simulation results above, Figure 1 shows that the tracking control objective of (6.1) is achievable. The output signal $x_{1,1}$ has realized tracking of the reference trajectory y_d. Corresponding to $g_{i,1} = \hat{x}_{i,1} - y_d$ in the article, it is proved that $g_{i,1}$ is convergent.

Figures 2 and 3 display the trajectory changes and their observed values of $x_{1,1}$ and $x_{1,2}$ during the two-step backstepping process. They indicate that the designed observer (2.17) can effectively observe the unknown states, and its observation error $x_{i,j} - \hat{x}_{i,j}$ converges. Figure 4 shows the error between the observed state $\hat{x}_{1,1}$ and the true state $x_{1,1}$.

Figures 5 and 6 present the variation diagrams of $||\hat{\omega}_{f1,1}||$ and $||\hat{\omega}_{f1,2}||$ during the two-step backstepping process. It can be seen that $||\hat{\omega}_{f1,1}||$ and $||\hat{\omega}_{f1,2}||$ tend to be stable. We can also prove their stability by constructing a Lyapunov function. Let $V = \hat{\omega}_{fi,j}^T \hat{\omega}_{fi,j}$, and the proof can be completed with the help of $\dot{\hat{\omega}}_{fi,j} = -\theta_{i,j}\psi_{fi,j}^T(\bar{x}_{i,j})\psi_{fi,j}(\bar{x}_{i,j})\hat{\omega}_{fi,j}$.

Figures 7 and 8 depict the trajectories of the norms of the evaluator-actor weight matrices over time, from which it can be observed that these trajectories tend to stabilize.

Figure 9 presents the optimal control input derived via the reinforcement learning algorithm. Subsequently, the critic network evaluates this input, while the actor network refines it accordingly on the basis of the critic's assessment, thereby ensuring that the control law aligns with the desired performance specifications, as illustrated in Figures 10 and 12. Figure 11 compares the trajectories of $\delta_{i,j}^x$ and $\hat{\delta}_{i,j}^x$, showing that the optimized $\hat{\delta}_{i,j}^x$ is more stable.

(2) Consider the system (2.1) as follows:

$$\begin{cases} \dot{x}_{2,1} = x_{2,2}, \\ \dot{x}_{2,2} = \frac{1}{M}(u_2 - \frac{mglsin(\chi)}{2}), \\ y_2 = x_{2,1}. \end{cases} \tag{6.2}$$

The system variables are consistent with those in reference [7, 33]. Among them, the parameters are defined as follows: M=1.2 kg·m^2, l=0.6 m, g=9.8 m/s^2, and m=0.08 kg. Select the the leader's output as y_d=0.6sin(1.2t) + 0.3cos(0.8t). Simultaneously select the initial values as $x_2=[x_{2,1}(0), x_{2,2}(0)]^T=[0.9, 0.8]^T$, and select the observed value of the initial value as $\hat{x}_{2,1}(0)$=0.8, $\hat{x}_{2,2}(0)$=0.7, $\hat{\omega}_{f2,1}=[0.8, 0.8, 0.8, 0.8, 0.8]^T$, $\hat{\omega}_{f2,2}=[0.8, 0.8, 0.8, 0.6, 0.6]^T$, $\hat{\omega}_{a1}(0)=[0.6, 0.6, 0.6, 0.6, 0.6]^T$, $\hat{\omega}_{a2}(0)=[0.6, 0.6, 0.6, 0.6, 0.6]^T$, $\hat{\omega}_{c1}(0) = [0.6, 0.6, 0.6, 0.6, 0.6]^T$ and $\hat{\omega}_{c2}(0)=[0.6, 0.6, 0.6, 0.6, 0.6]^T$. Here, $\hat{\omega}_{a2,1}(0) = [\hat{\omega}_{a1}(0), \hat{\omega}_{a1}(0)]^T$, $\hat{\omega}_{a2,2}(0) = [\hat{\omega}_{a2}(0), \hat{\omega}_{a2}(0)]^T$, $\hat{\omega}_{c2,1}(0) = [\hat{\omega}_{c1}(0), \hat{\omega}_{c1}(0)]^T$, and $\hat{\omega}_{c2,2}(0) = [\hat{\omega}_{c2}(0), \hat{\omega}_{c2}(0)]^T$. Select the control input parameters as $\eta_{2,2}$=30 and $\nu_{2,2}$=0.1. Select the parameters of the weight vector update law as $\theta_{2,1}$=10 and $\theta_{2,2}$=3. Select the parameters of the weight vector update law as $\beta_{a2,1}$=4, $\beta_{a2,2} = 4$, $\beta_{c2,1} = 8$, and $\beta_{c2,2} = 8$. The observer gains are set as $w_{2,1} = w_{2,2} = 50$.

Select the matrix P_2 as follows:

$$P_2 = \begin{bmatrix} 0.0371 & 0.0365 \\ 0.0365 & 4.0365 \end{bmatrix}.$$

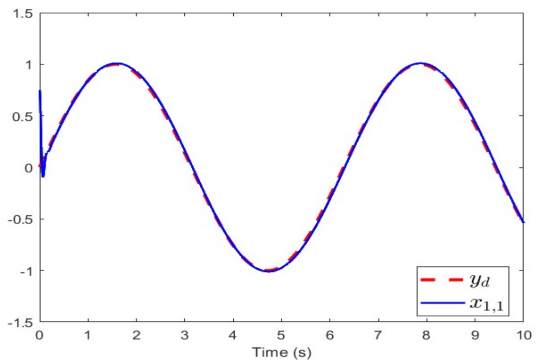

Figure 1. Tracking performance of (6.1).

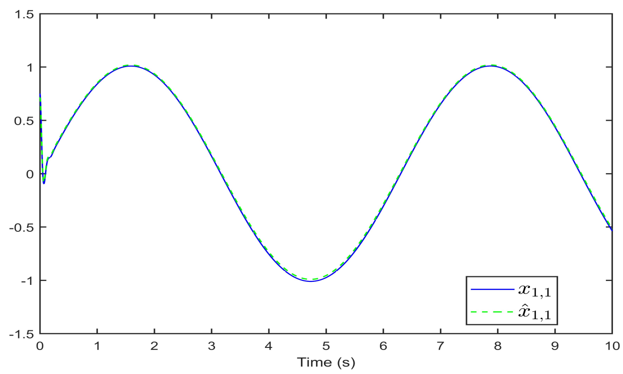

Figure 2. The curves of $x_{1,1}$ and $\hat{x}_{1,1}$.

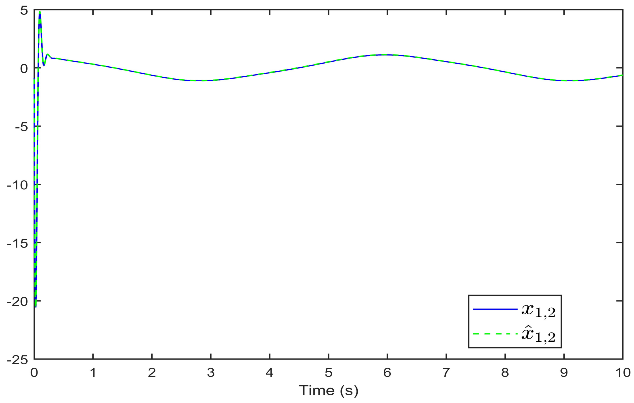

Figure 3. The curves of $x_{1,2}$ and $\hat{x}_{1,2}$.

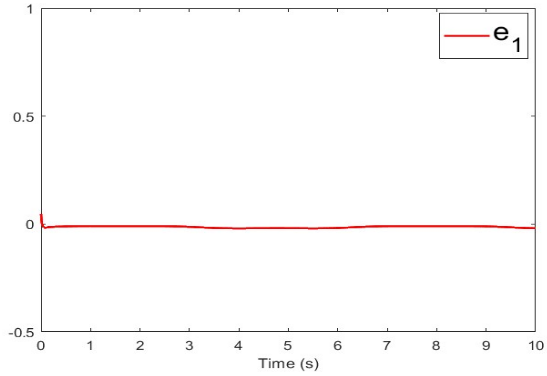

Figure 4. The observation error of system (6.1).

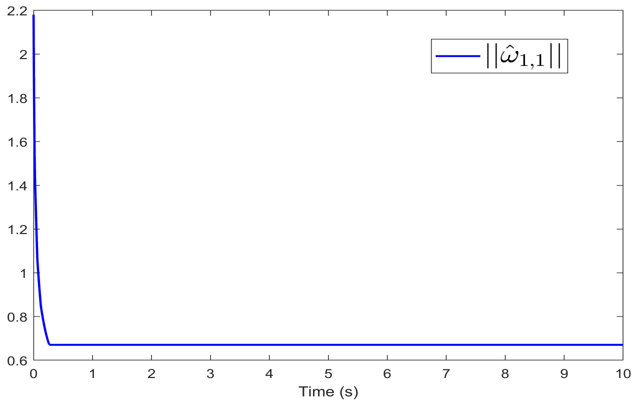

Figure 5. The trajectory variation diagram of $||\hat{\omega}_{1,1}||$.

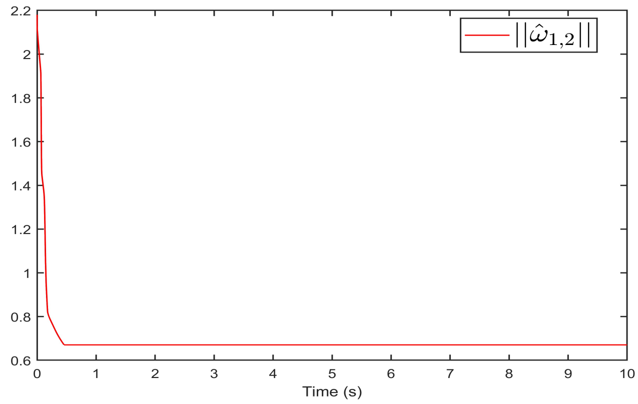

Figure 6. The trajectory variation diagram of $||\hat{\omega}_{1,2}||$.

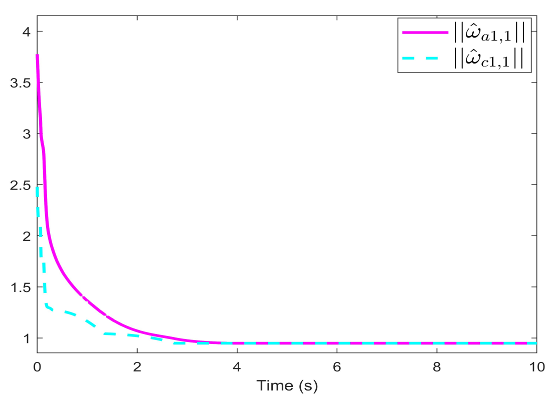

Figure 7. The trajectory diagram of $\hat{\omega}_{a1,1}$ and $\hat{\omega}_{c1,1}$.

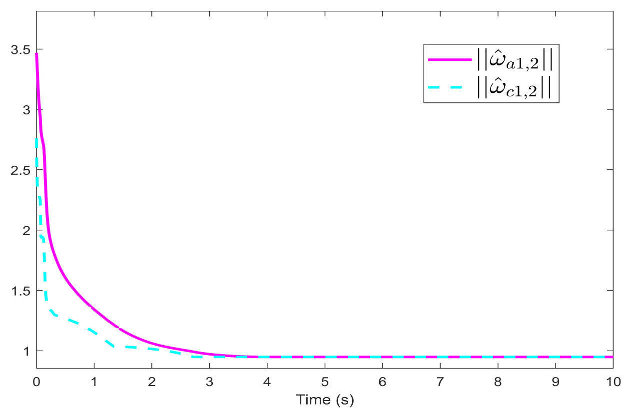

Figure 8. The trajectory diagram of $\hat{\omega}_{a1,2}$ and $\hat{\omega}_{c1,2}$.

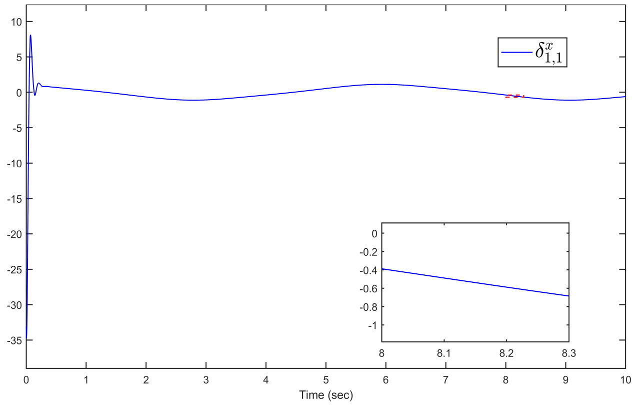

Figure 9. The control input $\delta^x_{1,1}$.

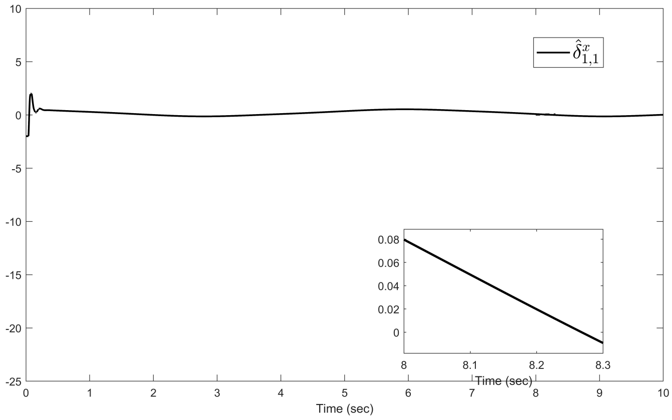

Figure 10. The control input $\hat{\delta}^x_{1,1}$.

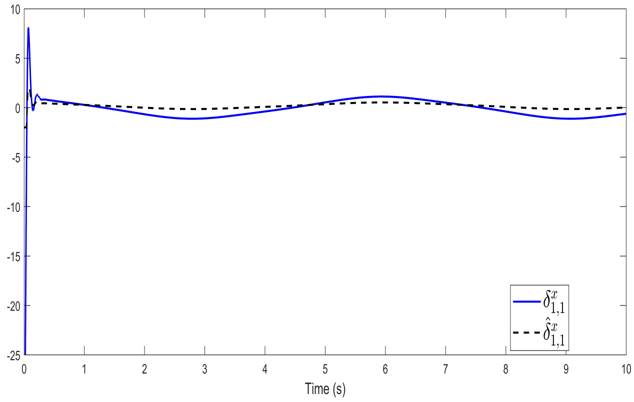

Figure 11. Difference between $\delta^x_{1,1}$ and $\hat{\delta}^x_{1,1}$.

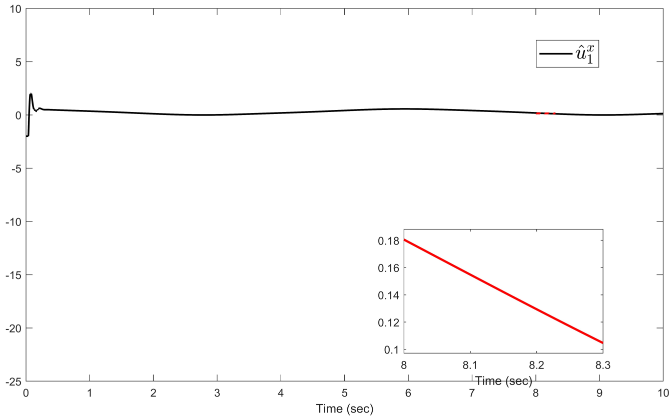

Figure 12. The control input $\hat{u}^x_1$.

It can be seen from Figure 13 that the output signal $x_{2,1}$ can track the output signal y_d of the leader. On one hand, this is because $y_i = x_{i,1}$ in the system (2.1); on the other hand, $g_{i,1} = \hat{x}_{i,1} - y_d$ converges, and the observed value $\hat{x}_{i,1}$ can replace the true value $x_{i,1}$, enabling the trajectory of $x_{i,1}$ to track the trajectory of y_d. Figures 14 and 15 show the trajectories of the states $x_{2,1}$ and $x_{2,2}$ and their observations. Figures 16 and 17 show the time-varying trajectories of the weight vector norms $||\omega_{fi,j}||$ of the unknown functions $f_{i,j}(x_{i,j}(t))$, all of which tend to stabliize in a short time. Similarly, due to $\dot{\hat{\omega}}_{fi,j} = -\theta_{i,j}\psi^T_{fi,j}(\bar{x}_{i,j})\psi_{fi,j}(\bar{x}_{i,j})\hat{\omega}_{fi,j}$, we can also prove their convergence through the Lyapunov stability theory. Figure 18 shows the optimal control input derived using the RL algorithm, while Figure 19 reflects the optimal control input adjusted on the basis of the evaluation of the actor–critic network.

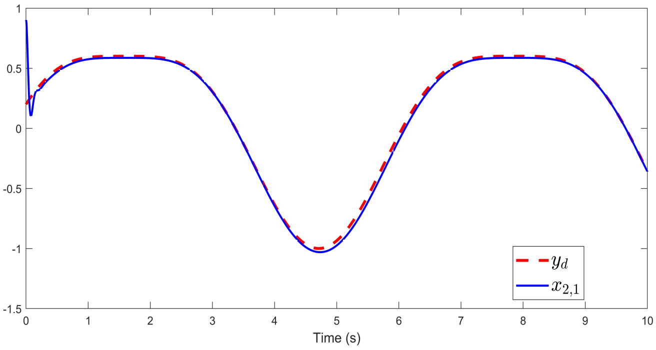

Figure 13. Tracking performance of (6.2).

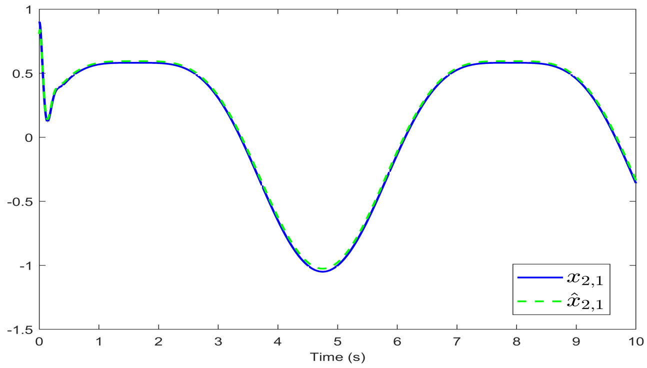

Figure 14. The trajectory of $x_{2,1}$ and $\hat{x}_{2,1}$.

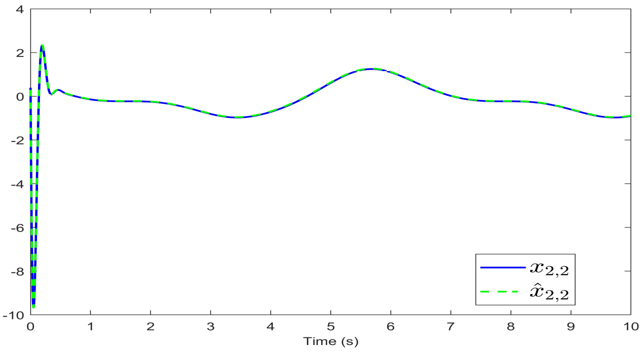

Figure 15. The trajectory of $x_{2,2}$ and $\hat{x}_{2,2}$.

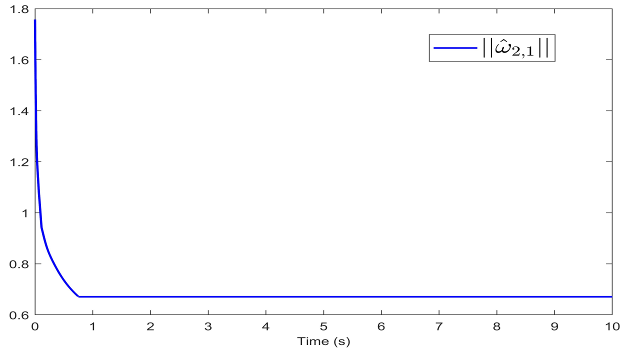

Figure 16. The trajectory variation diagram of $||\hat{\omega}._{2,1}||$.

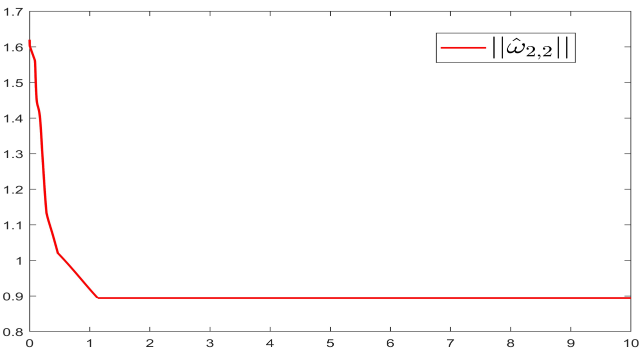

Figure 17. The trajectory variation diagram of $||\hat{\omega}_{2,2}||$.

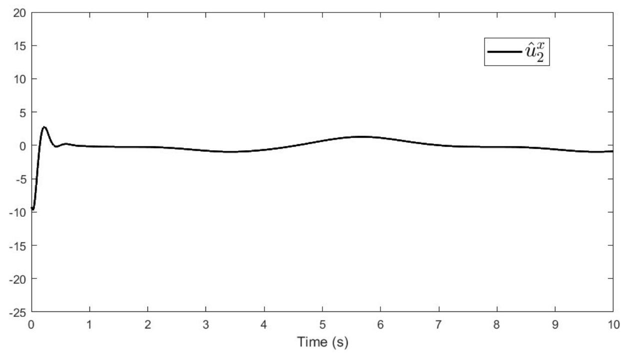

Figure 18. The control input $\hat{u}_2^x$.

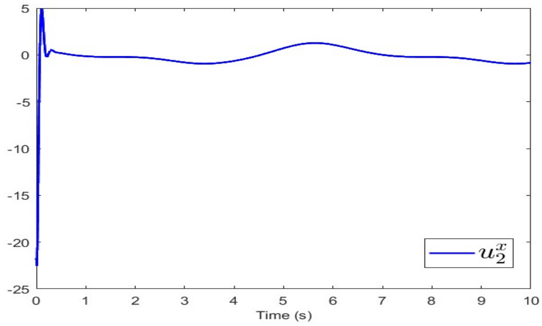

Figure 19. The control input u_2^x.

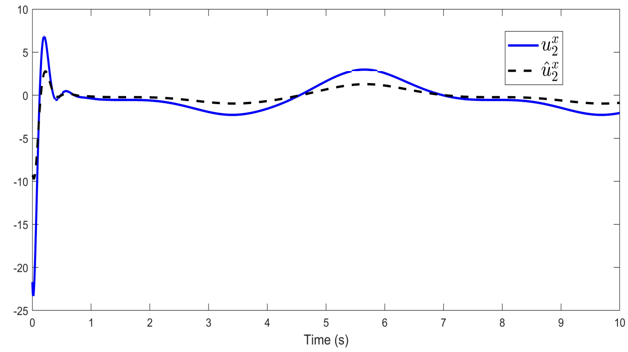

Figure 20. Difference between u_2^x and $\hat{u}_2^x$.

6.2. Comparative analysis of the relevant figures

To more fully demonstrate the superiority of the fixed-time consensus algorithm and the advantages of the proposed method, this paper takes the weights of the critic–actor network as the core research object and conducts a comparative analysis with the finite-time consensus algorithm (using the control method in this paper combined with the data from [29]) through a visual comparison.

Remark 6.1. *Figures 21 and 23 present the critic–actor network weight curves obtained by the method proposed in this paper; in contrast, Figures 22 and 24 show the weight curves simulated using the data from [29]. The curves in this work satisfy boundedness, stability, and monotonicity, whereas those in [29] only guarantee boundedness but fail to ensure stability and monotonicity. It is worth noting that the monotonicity and boundedness of* $||\hat{\omega}_{ai,j}||$ *and* $||\hat{\omega}_{ci,j}||$ *are necessary conditions for obtaining the high-order term of the Lyapunov function* (V^q)*. The upper bound obtained in this paper is the initial values* $||\hat{\omega}_{ai,j}(0)||$ *and* $||\hat{\omega}_{ci,j}(0)||$*. However, since the method in [29] cannot obtain accurate upper bounds, it can only construct the first-order term V of the Lyapunov function required for finite-time stability, and thus cannot achieve fixed-time convergence.*

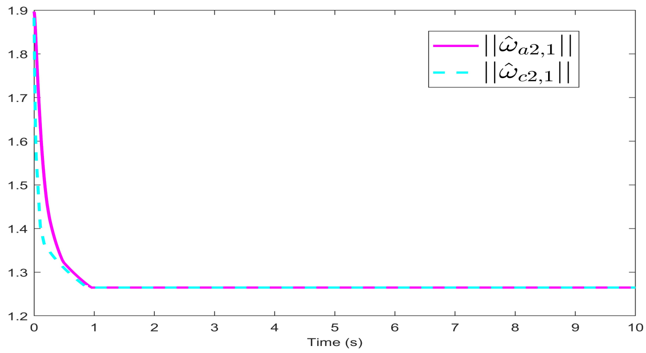

Figure 21. Norm of the critic–actor network weights.

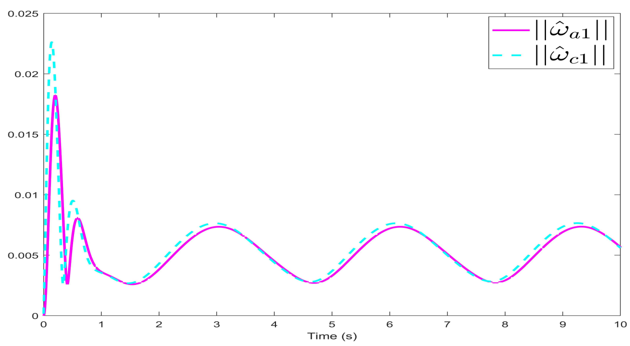

Figure 22. The norm of the critic–actor network weights obtained through simulation.

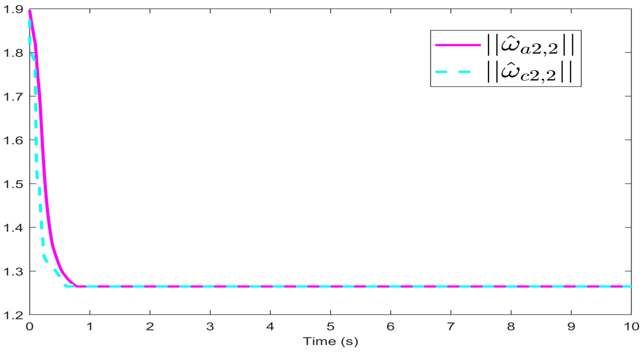

Figure 23. Norm of the critic–actor network weights.

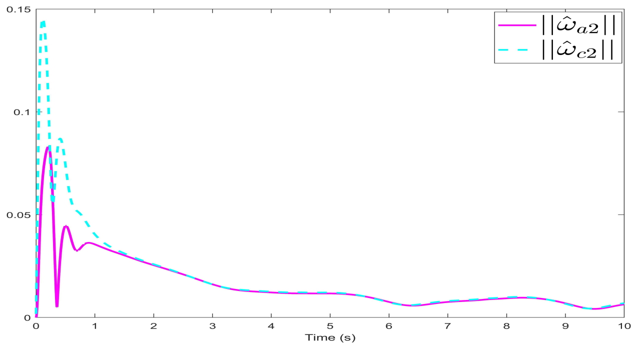

Figure 24. The norm of the critic–actor network weights obtained through simulation.

7. Conclusions

This paper studies the fixed-time optimal consensus of strict-feedback multi-agent systems under RL. The backstepping method and RL algorithm are applied to the observer–critic–actor framework to derive the control input of the system. Through online evaluation and adjustment of the control input, the optimal control input is obtained. Additionally, a Lyapunov function is constructed to achieve the fixed-time consensus of the system. The novelty of this paper lies in the fact that the dimensions of the critic–actor network are determined by introducing unknown nonlinear functions. Furthermore, while performing scalar calculations, the upper bounds of the weights of the critic–actor network are obtained, thereby solving the problem of the quadratic term (V^2) of the Lyapunov function required for achieving fixed-time stability. Future research will explore higher-order Lyapunov function terms (V^q, $q > 2$).

Author contributions

Kaile Zhang: Writing–original draft, validation; Zhanheng Chen: Writing–review and editing, supervision; Zhiyong Yu: Conceptualization, supervision, writing–review and editing; Haijun Jiang: Supervision, conceptualization, formal analysis, resources. All authors have read and approved the final version of the manuscript for publication.

Use of Genertive-AI tools declaration

The authors declare that they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This research was supported by the Natural Science Foundation of Xinjiang Uygur Autonomous Region (2024D01C196), the National Natural Science Foundation of China (62163035), the special fund project of the central government to guide local science and technology development

(ZYYD2022A05), and the Key Program of Yili Normal University (2024YSZD001).

Conflict of interest

The authors declare that there are no conflicts of interest.

References

1. J. Zhang, H. Xia, G. Ma, Fixed-time coordinated attitude control for spacecraft formation flying via event-triggered and delayed communication, *Adv. Space Res.,* **73** (2024), 6140–6146. https://doi.org/10.1016/j.asr.2024.03.020
2. X. Dong, H. Zhang, X. Xie, Z. Ming, Data-driven distributed H_∞ current sharing consensus optimal control of DC microgrids via reinforcement learning, *Adv. Space Res.,* **71** (2024), 2824–2834. https://doi.org/10.1109/TCSI.2024.3366942
3. A. Coppola, D. G. Lui, A. Petrillo, S. Santini, Cooperative driving of heterogeneous uncertain nonlinear connected and autonomous vehicles via distributed switching robust pid-like control, *Inf. Sci.,* **625** (2023), 277–298. https://doi.org/10.1016/j.ins.2023.01.045
4. M. H. Moradi, S. Razini, S. M. Hosseinian, State of art of multiagent systems in power engineering: A review, *Renew. Sust. Energ. Rev.,* **58** (2016), 814–824. https://doi.org/10.1016/j.rser.2015.12.339
5. W. Wu, Y. Li, S. Tong, Neural network output-feedback consensus fault-tolerant control for nonlinear multiagent systems with intermittent actuator faults, *IEEE T. Neural Networ.,* **34** (2024), 4728–4740. https://doi.org/10.1109/TNNLS.2021.3117364
6. W. Su, C. Mu, S. Zhu, Event-triggered leader-follower bipartite consensus control for nonlinear multi-agent systems under DoS attacks, *Sci. China Inf. Sci.,* **68** (2024), 132206. https://doi.org/10.1007/s11432-024-4148-7
7. X. Wu, X. He, L. Wu, Finite-time adaptive fault-tolerant control for a class of nonlinear systems with actuator faults, *J. Control Autom. Electr. Syst.*, **31** (2020), 1075–1086. https://doi.org/10.1007/s40313-020-00618-4
8. Y. Li, G. Lu, K. Li, Fuzzy adaptive event-triggered consensus control for nonlinear multiagent systems with output constraints and DoS attacks, *IEEE T. Cybernetics*, **55** (2025), 2–13. https://doi.org/10.1109/TCYB.2024.3456821
9. T. Yu, J. Cao, K. Lu, Finite-time synchronization control of networked chaotic complex-valued systems with adaptive coupling, *IEEE T. Netw. Sci. Eng.*, **9** (2022), 2510–2520. https://doi.org/10.1109/TNSE.2022.3164773
10. X. Chen, J. Sun, Y. Wang, Finite-time control for attitude tracking of spacecraft with actuator faults, *Robotics*, **10** (2020), 128.
11. C. Huang, Z. Liu, C. L. P. Chen, Y. Zhang, Adaptive fixed-time neural control for uncertain nonlinear multiagent systems, *IEEE T. Neural Networ.*, **34** (2022), 10346–10358. https://doi.org/10.1109/TNNLS.2022.3165836
12. F. Wang, G. Lai, Fixed-time control design for nonlinear uncertain systems via adaptive method, *Syst. Control Lett.*, **140** (2020), 104704. https://doi.org/10.1016/j.sysconle.2020.104704

13. A. Polyakov, Nonlinear feedback design for fixed-time stabilization of linear control systems, *IEEE Trans. Automat. Contr.*, **57** (2012), 2106–2110. https://doi.org/10.1109/TAC.2011.2179869
14. P. Wang, C. Yu, M. Lv, J. Cao, Adaptive fixed-time optimal formation control for uncertain nonlinear multiagent systems using reinforcement learning, *IEEE T. Netw. Sci. Eng.*, **12** (2024), 1729–1743. https://doi.org/10.1109/TNSE.2023.3330266
15. S. Du, J. Han, H. Y. Sun, H. G. Han, J. Qiao, Fully distributed practical fixed-time consensus of second-order nonlinear multiagent systems under actuator faults, *J. Franklin I.*, **362** (2025), 108009. https://doi.org/10.1016/j.jfranklin.2025.108009
16. S. Wang, X. Zhan, J. Wu, B. Wu, L. Chen, Fixed-time consensus algorithm for second-order multi-agent systems with actuator faults based on non-singular terminal sliding mode, *Circ. Syst. Signal Pr.*, **44** (2025), 1–22. https://doi.org/10.1007/s00034-025-03039-3
17. L. Shi, W. Guo, L. Wang, W. Sun, H. Jahanshahi, Fixed-time consensus control for nonlinear multi-agent systems via fully distributed edge-based adaptive protocols, *Int. J. Control*, **97** (2024), 1576–1585. https://doi.org/10.1080/00207179.2023.2215068
18. A. Liu, S. He, H. Yang, Observer-based double event-triggered fast finite-time Cconsensus control for heterogeneous nonlinear multi-agent systems, *Nonlinear Dyn.*, **113** (2025), 1–19. https://doi.org/10.1007/s11071-025-11423-6
19. W. T. Miller, R. S. Sutton, P. J. Werbos, A menu of designs for reinforcement learning over time, *IEEE T. Neural Networ.*, **3** (1995), 67–95.
20. T. Chen, J. Yuan, H. Yang, Event-triggered adaptive neural network backstepping sliding mode control of fractional-order multi-agent systems with input delay, *J. Vib. Control*, **28** (2022), 1008–1021.
21. Y. Wu, G. Zhang, L. B. Wu, Event-triggered adaptive fault-tolerant control for non-affine uncertain systems with output tracking errors constraints, *IEEE T. Fuzzy Syst.*, **30** (2021), 1750–1761. https://doi.org/10.1109/TFUZZ.2021.3065538
22. J. Wu, F. He, H. Shen, S. Ding, Z. G. Wu, Adaptive NN fixed-time fault-tolerant control for uncertain stochastic system with deferred output constraint via self-triggered mechanism, *IEEE T. Cybernetics*, **53** (2022), 5892–5903. https://doi.org/10.1109/TCYB.2022.3205765
23. Y. Wang, Y. Yang, L. Wu, Adaptive fault-tolerant consensus control of multi-agent systems with event-triggered inputs, *Inf. Sci.*, **650** (2023), 119594. https://doi.org/10.1016/j.ins.2023.119594
24. J. Wu, H. Lu, W. Wang, *Adaptive deep neural network optimized backstepping control for a class of nonlinear strict-feedback systems*, In: 2024 12th International Conference on Intelligent Control and Information Processing (ICICIP), 2024, 172–181. https://doi.org/10.1109/ICICIP60808.2024.10477826
25. B. Zhu, H. Liang, B. Niu, H. Wang, N. Zhao, X. Zhao, Observer-based reinforcement learning for optimal fault-tolerant consensus control of nonlinear multiagent systems via a dynamic event-triggered mechanism, *Inf. Sci.*, **689** (2025), 121350. https://doi.org/10.1016/j.ins.2024.121350
26. D. Li, J. Dong, Fuzzy control based on reinforcement learning and subsystem error derivatives for strict-feedback systems with an observer, *IEEE T. Fuzzy Syst.*, **31** (2022), 2509–2521. https://doi.org/10.1109/TFUZZ.2022.3227993

27. J. Liu, Y. Dong, G. Zhou, Security consensus control for multi-agent systems under DoS attacks via reinforcement learning method, *J. Franklin I.*, **361** (2024), 164–176. https://doi.org/10.1016/j.jfranklin.2023.11.032

28. X. Song, P. Sun, C. K. Ahn, S. Song, Reinforcement learning-based event-triggered predefined-time optimal fuzzy control for nonlinear constrained systems, *IEEE Trans. Fuzzy Syst.*, **32** (2024), 4646–4659. https://doi.org/10.1109/TFUZZ.2024.3408847

29. Y. Sun, G. Lu, K. Peng, L. Wu, Finite-time adaptive optimal control of uncertain strict-feedback nonlinear systems based on fuzzy observer and reinforcement learning, *Int. J. Syst. Sci.*, **55** (2023), 1553–1570. https://doi.org/10.1080/00207721.2024.2312882

30. W. Qi, C. Zhang, G. Zong, S. F. Su, M. Chadli, Finite-time event-triggered stabilization for discrete-time fuzzy markov jump singularly perturbed systems, *IEEE T. Cybernetics*, **53** (2023), 4511–4520. https://doi.org/10.1109/TCYB.2022.3207430

31. B. Zhu, N. Xu, G. Zong, X. Zhao, Adaptive optimized backstepping tracking control for full-state constrained nonlinear strict-feedback systems without using barrier Lyapunov function method, *Optim. Contr. Appl. Met.*, **45** (2024), 2051–2075. https://doi.org/10.1002/oca.3136

32. R. Hao, H. Wang, W. Zheng, Finite-time dynamic event-triggered full-state constraints consensus control for multiagent systems with switching topologies and mismatched disturbances, *J. Intell. Robot. Syst.*, **110** (2024), 147. https://doi.org/10.1007/s10846-024-02168-2

33. Y. J. Liu, S. Tong, Barrier Lyapunov functions-based adaptive control for a class of nonlinear pure-feedback systems with full state constraints, *Automatica*, **64** (2016), 70–75. https://doi.org/10.1016/j.automatica.2015.10.034