Optimized Multi-Agent Formation Control Based on an Identifier–Actor–Critic Reinforcement Learning Algorithm

Guoxing Wen ⃝, C. L. Philip Chen ⃝, *Fellow, IEEE*, Jun Feng, and Ning Zhou

Abstract—The paper proposes an optimized leader–follower formation control for the multi-agent systems with unknown nonlinear dynamics. Usually, optimal control is designed based on the solution of the Hamilton–Jacobi–Bellman equation, but it is very difficult to solve the equation because of the unknown dynamic and inherent nonlinearity. Specifically, to multi-agent systems, it will become more complicated owing to the state coupling problem in control design. In order to achieve the optimized control, the reinforcement learning algorithm of the identifier–actor–critic architecture is implemented based on fuzzy logic system (FLS) approximators. The identifier is designed for estimating the unknown multi-agent dynamics; the actor and critic FLSs are constructed for executing control behavior and evaluating control performance, respectively. According to Lyapunov stability theory, it is proven that the desired optimizing performance can be arrived. Finally, a simulation example is carried out to further demonstrate the effectiveness of the proposed control approach.

Index Terms—Fuzzy logic systems (FLSs), identifier–actor–critic architecture, multi-agent formation, optimized formation control, reinforcement learning (RL).

G. X. Wen is with the College of Science, Binzhou University, Binzhou 256600, China (e-mail: gxwen@live.cn).

C. L. P. Chen is with the Department of Computer and Information Science, Faculty of Science and Technology, University of Macau, Macau 99999, China, with Dalian Maritime University, Dalian 116026, China, and also with State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100080, China (e-mail: philip.chen@ieee.org).

J. Feng is with the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210000, China, and also with the Department of Information Engineering, Binzhou University, Binzhou 256600, China (e-mail: fengjunImportant@163.com).

N. Zhou is with the Faculty of Science and Engineering, University of Groningen, Groningen 9747 AG, The Netherlands, and also with the College of Computer and Information Sciences, Fujian Agriculture and Forestry University, Fuzhou 350002, China (e-mail: zhouning2010@gmail.com).

I. INTRODUCTION

IN THE multi-agent cooperation community, formation control is one of the most interesting and attractive research topics because of its broad applications, such as cooperative control of unmanned aerial vehicles, satellite clusters, autonomous underwater vehicles, and mobile sensor networks. In brief, formation control is to design the appropriate protocol or algorithm such that the multi-agent system arrives and maintains a predefined geometrical shape, for example, a chain or wedge. In the recent decades, formation control has been well developed, and several published results receive the considerable and increasing attention, such as leader–follower [1], behavior [2], virtual structure [3], and potential function based approaches [4], where the leader–follower approach is the most popular one due to its simplicity and scalability. The basic idea is that a leader is designed as a reference for the agent group, and all agents as followers are controlled to maintain the desired separation and relative bearing with the leader. The main advantage is that group behavior is specified by a single quantity (the leader's motion).

Ever since optimal control, which means that cost function is minimized, was formally developed about five decades ago by Bellman [5] and Pontryagin [6], optimization became a fundamental design idea and principle in modern control theory. In recent years, the optimal problem has been addressed in formation control of multi-agent systems, and several approaches have been published [7]–[9]. In [7], the finite-time optimal formation problem of multi-agent systems on the Lie group SE(3) is investigated. In [8], the finite time optimal formation is applied to multivehicle systems. In [9], the centralized optimal multi-agent coordination problem under tree formation constraints is studied. These published optimal formation methods are achieved based on the solution of the Hamilton-Jacobi-Bellman (HJB) or Hamiltonian equation. In practice, the HJB equation is solved diffreultly by analytical approaches owing to the inherent nonlinearities and unknown dynamics.

In order to overcome the difficulty coming from solving the HJB equation, a reinforcement learning (RL)-based function approximation strategy is usually considered. The basic idea is that appropriate actions are taken by evaluating feedback from environment [10]. One of the most popular means to perform RL algorithms is the actor–critic architecture, where the actor performs certain actions by interacting with environment and the critic evaluates the actions and gives feedback to the actor [11].

However, most of the RL-based optimal approaches require complete knowledge of system dynamics, and it is difficult to be satisfied for practical situations. In order to release the strict requirement, an effective solution is the identifier–actor–critic method because the unknown dynamics are estimated by the identifier for RL [12].

It is well known that fuzzy logical systems (FLSs) have excellent approximation ability, which can approximate any continuous function to the desired accuracy over a compact set. In the recent years, many frequently used control techniques have been well developed based on the FLS approximator, such as backstepping, optimizer, small-gain approach, and dead-zone control [13]–[16], and widely applied to various nonlinear systems, such as [17]–[22]. However, a common challenge and difficulty in adaptive fuzzy control is the stability proof because there possibly exists the undesirable drift in the online learning. Recently, several stability analysis approaches are published to gain the extensive attention [23]–[25], they are the effective ways for solving the difficulty. Nevertheless, for multi-agent system control, stability analysis becomes more challenging and difficult owing to the state coupling in the control design. To the optimized formation control, stability analysis is turned into a very complex and intractability problem because RL is performed by online training both critic and actor simultaneously.

Motivated by the above-mentioned discussion, in this paper, the RL algorithm of the identifier–actor–critic architecture is utilized for the optimized formation control. Based on FLS approximations of the unknown nonlinear dynamic and optimal value functions, the identifier, actor, and critic are constructed, where the online learning for them is continuous and simultaneous. The main contributions are listed in the following.

1) The optimized formation control approach can efficiently solve the tracking problem by segmenting an error term from the optimal value function. Owing to the difficulty in the convergence analysis of tracking errors, existing optimization control methods rarely involve the tracking problem. The proposed optimization strategy can well carry out tracking control; therefore, it can guarantee that the leader–follower formation control is fulfilled.

2) The RL of the identifier–actor–critic architecture is applied to multi-agent control so that the excellent control performance can be guaranteed. Most of the existing RLs are designed based on a common assumption that the system dynamics are completely known, such as [26] and [27]. However, this assumption is impractical or very strict for many practical situations. The proposed RL algorithm can release the strict assumption because the adaptive identifier is employed to estimate the system uncertainties, it can meet the practical requirements for real-world engineering.

3) The strict proofs for the stability and convergence analyses are given. In most of the existing RL control literature, Lyapunov function for stability analysis is designed to contain the infinite horizon value function, such as [12] and [28]. Because the function's derivative is negative, it cannot guarantee that the strict analyses are performed for stability and convergence.

For convenience, the following notations are used throughout the paper.

1) R represents the real number; R^n denotes the real n-dimensional vector space; $R^{n \times m}$ is the $n \times m$-dimensional matrix space; and I_n is the $n \times n$ identity matrix.

2) $|\cdot|$ denotes the absolute value; $\|\cdot\|$ represents the 2-norm; and Ω represents the set.

3) T is the transposition symbol; and $\otimes$ denotes the Kronecker product.

II. PRELIMINARIES

A. Fuzzy Logic Systems

It has been proven that FLSs have the universal approximation and learning abilities. A FLS is composed of four parts, which are the knowledge base, fuzzifier, fuzzy inference engine, and defuzzifier.

The knowledge base is a collection of fuzzy If-Then rules described in the following:

$$R_j : \text{If } x_1 \text{ is } F_1^j \text{ and } x_2 \text{ is } F_2^j \ldots \text{ and } x_n \text{ is } F_n^j$$
$$\text{Then } y \text{ is } G^j, \ j = 1, 2, \ldots, N$$

where $x = [x_1, \ldots, x_n]^T$ is the input; y is the output; F_i^j and G^j are the fuzzy sets associated with fuzzy membership functions $\mu_{F_i^j}(x_i) \in R$ and $\mu_{G^j}(y) \in R$, respectively; and N is the number of rules.

The singleton fuzzifier, product inference engine, and center-average defuzzifier are defined as

$$y(x) = \frac{\sum_{j=1}^{N} \left(\theta_j \prod_{i=1}^{n} \mu_{F_i^j}(x_i) \right)}{\sum_{j=1}^{N} \left(\prod_{i=1}^{n} \mu_{F_i^j}(x_i) \right)} \tag{1}$$

where $\theta_j = \max\limits_{y \in R} \mu_{G^j}(y)$.

Define the fuzzy basis function as

$$\varphi_j(x) = \frac{\prod_{i=1}^{n} \mu_{F_i^j}(x_i)}{\sum_{j=1}^{N} \left(\prod_{i=1}^{n} \mu_{F_i^j}(x_i) \right)} \tag{2}$$

the FLS (1) can be re-expressed as

$$y(x) = \Theta^T \varphi(x) \tag{3}$$

where $\Theta = [\theta_1, \ldots, \theta_N]^T$ is viewed as the adjustable parameter vector and $\varphi(x) = [\varphi_1(x), \ldots, \varphi_N(x)]^T$ is the fuzzy basis function vector.

It has been proven that the FLS can uniformly approximate any continuous nonlinear function to the desired accuracy over a compact set. This property is described by the following lemma.

Lemma 1: [29] Any real continuous function $h(x) \in R$ is well defined on a compact set $\Omega_h \in R^n$, there exists the FLS described by (3) such that

$$\sup_{x \in \Omega_h} |h(x) - y(x)| < \varepsilon$$

where $\varepsilon > 0$ is an arbtrary positive number.

According to Lemma 1, for any continuous vector-valued function $f(x) = [f_1(x), \ldots, f_m(x)]^T \in R^m$ defined on the compact set $\Omega_f \in R^m$, there exists an optimal parameter matrix $\Theta_f^* = [\Theta_{f1}^*, \ldots, \Theta_{fm}^*] \in R^{N \times m}$ such that

$$f(x) = \Theta_f^{*T} \varphi(x) + \varepsilon_f(x) \tag{4}$$

where $\varepsilon_f(x) \in R^m$ is the approximation error satisfying $\|\varepsilon_f(x)\| \leq \delta$, δ is a positive constant. The optimal parameter vector Θ_f^* is defined as

$$\Theta_f^* := \arg \min_{\Theta \in R^{N \times m}} \left\{ \sup_{x \in \Omega_f} \left\| f(x) - \Theta_f^T \varphi(x) \right\| \right\} \tag{5}$$

where $\Theta_f = [\Theta_{f1}, \ldots, \Theta_{fm}] \in R^{N \times m}$ is the adjustable parameter matrix. It should be mentioned that Θ_f^* needs to be estimated because it is an "artificial" quantity just for analysis purposes.

B. Algebraic Graph Theory

The interconnection topology of a multi-agent system can be depicted by a graph $G = (\Upsilon, \Xi, A)$, where $\Upsilon = \{v_1, v_2, \ldots, v_n\}$, $\Xi \subseteq \Upsilon \times \Upsilon$ and $A = [a_{ij}]$ are the node set, edge set, and adjacency matrix, respectively. Let $\xi_{ij} = (v_i, v_j)$ denote the edge connecting both agents i and j, then $\xi_{ij} \in \Xi$ if and only if there is an information flow from agent j to agent i. Agent j is called as a neighbor of agent i if $\xi_{ij} \in \Xi$, and the neighbor set of agent i is denoted by $\Lambda_i = \{v_j \| (v_i, v_j) \in \Xi\}$. The adjacency element a_{ij} denotes the communication weight corresponding to the edge ξ_{ij}, which satisfies $\xi_{ij} \in \Xi \Leftrightarrow a_{ij} = 1$ and otherwise $a_{ij} = 0$. A graph G is called undirected if $a_{ij} = a_{ji}$. An undirected graph is called connected if any a pair of distinct nodes can be connected by an undirected path. The Laplacian matrix $L = [l_{ij}] \subset R^{n \times n}$ of the weight graph G is defined as

$$L = D - A \tag{6}$$

where $d = \text{diag}\{d_1, \ldots, d_n\}$, $d_i = \sum_{j=1}^n a_{ij}$.

Let b_i denote the connection weight between agent i and the leader. If there is the information communication between agent i and the leader, then $b_i = 1$, otherwise $b_i = 0$. It is assumed that at least one agent connects with the leader, i.e., $b_1 + b_2 + \cdots + b_n > 0$.

C. Supporting Lemmas

Lemma 2: [30] An undirected graph G is connected if and only if its Laplacian is irreducible.

Lemma 3: [30] Let $Q = [q_{ij}] \in R^{n \times n}$ be an irreducible matrix such that $q_{ij} = q_{ji} \leq 0$ for $i \neq j$ and $q_{ii} = -\sum_{j=1}^n q_{ij}$ for $i = 1, 2, \ldots, n$. Then all eigenvalues of the matrix

$$\begin{bmatrix} q_{11} + \bar{q}_1 & \cdots & q_{1n} \\ \vdots & \ddots & \vdots \\ q_{n1} & \cdots & q_{nn} + \bar{q}_n \end{bmatrix}$$

are positive, where $\bar{q}_1, \bar{q}_2, \ldots, \bar{q}_n$ are non-negative constants satisfying $\bar{q}_1 + \bar{q}_2 + \cdots + \bar{q}_n > 0$.

Lemma 4: [30] Let $\Phi(t) \in R$ be a continuous positive function with bounded initial value $\Phi(0)$. If $\dot{\Phi}(t) \leq -\alpha \Phi(t) + \beta$

is held, where α and β are positive constants, then there is the following result:

$$\Phi(t) \leq e^{-\alpha t} \Phi(0) + \frac{\beta}{\alpha} \left(1 - e^{-\alpha t}\right). \tag{7}$$

III. MAIN RESULTS

A. Problem Formulation

Consider the multi-agent system modeled in the following:

$$\dot{x}_i(t) = f_i(x_i(t)) + u_i, \quad i = 1, \ldots, n \tag{8}$$

where $x_i(t) \in R^m$ is the state; $u_i \in R^m$ is the control input; and $f_i(\cdot): R^m \to R^m$ with $f_i(0) = 0_n$ is the unknown nonlinear continuous vector-value function. These terms $f_i(x_i) + u_i$, $i = 1, 2, \ldots, n$, are assumed Lipschitz continuous on the set containing origin so that the solution of differential equation (8) is unique for any bounded initial state $x_i(0)$. The system (8) is assumed stabilizable, i.e., there exists the continuous control u_i such that the system is asymptotically stable. The communication graph G is assumed to be an undirected connected graph.

Let $x_d(t), \dot{x}_d(t) \in R^m$ denote the desired trajectory and velocity of the formation movement, which are assumed known and bounded. Define the tracking error variable for agent i as

$$z_i(t) = x_i(t) - x_d(t) - \eta_i, \quad i = 1, 2, \ldots, n \tag{9}$$

where $\eta_i = [\eta_{i1}, \eta_{i2}, \ldots, \eta_{im}]^T$ is the relative position vector between agent i and the leader, which depicts the predefined formation pattern.

Definition 1: [31] The multi-agent system (8) is said to achieve the desired formation if its solutions satisfy

$$\lim_{t \to \infty} \|x_i(t) - x_d(t) - \eta_i\| = 0, \quad i = 1, \ldots, n$$

for the bounded initial conditions.

Based on (8), the following error dynamic can be yielded:

$$\dot{z}_i(t) = f_i(x_i) - \dot{x}_d(t) + u_i, \quad i = 1, \ldots, n. \tag{10}$$

Define the formation errors as

$$e_i(t) = \sum_{j \in \Lambda_i} a_{ij}(x_i(t) - \eta_i - x_j(t) + \eta_j)$$
$$+ b_i(x_i(t) - x_d(t) - \eta_i), \quad i = 1, \ldots, n \tag{11}$$

where a_{ij} is the ith row and jth column element of adjacency matrix A; and b_i is the connection weight between agent i and the leader. Inserting (9) into (11), the following equation can be yielded:

$$e_i(t) = \sum_{j \in \Lambda_i} a_{ij}(z_i - z_j) + b_i z_i, \quad i = 1, \ldots, n. \tag{12}$$

Based on the multi-agent dynamic (8), time derivative of the formation error is

$$\dot{e}_i(t) = c_i f_i(x_i) + c_i u_i - b_i \dot{x}_d(t) - \sum_{j \in \Lambda_i} a_{ij} \dot{x}_j(t) \tag{13}$$

where $c_i = \sum_{j \in \Lambda_i} a_{ij} + b_i$.

Define the infinite horizon value function as

$$V(e(t)) = \int_t^\infty r(e(\tau), u(e)) d\tau \tag{14}$$

where $r(e, u) = e^T(t)e(t) + u^T(C \otimes I_m)u = z^T(t)(\tilde{L}^T \tilde{L} \otimes I_m)z(t) + u^T(C \otimes I_m)u$ is the cost function, where $e^T(t) = [e_1^T, \ldots, e_n^T]$; $u = [u_1^T, \ldots, u_n^T]^T$; $z = [z_1^T, \ldots, z_n^T]^T$; $C = \text{diag}\{c_1, \ldots, c_n\}$; and $\tilde{L} = L + B$. It should be mentioned that $\tilde{L}$ is a positive definite matrix in accordance with Lemma 3.

Let $r_i(e_i, u_i) = e_i^T e_i + c_i u_i^T u_i$ and $V_i(e_i) = \int_t^\infty r_i(e_i(\tau), u_i(e_i)) d\tau$, the value function (14) can be re-expressed as

$$V(e) = \sum_{i=1}^n V_i(e_i) = \sum_{i=1}^n \int_t^\infty r_i(e_i(\tau), u_i(e_i)) d\tau. \quad (15)$$

Definition 2: [32] The multi-agent formation control u_i, $i = 1, \ldots, n$, is said to be admissible associating with (10) on a set $\bar{\Omega}$, which is denoted by $u_{i=1,\cdots,n} \in \Psi(\bar{\Omega})$, if u_i, $i = 1, \ldots, n$, is continuous with $u_i(0) = 0$, u_i stabilizes (10) and $V(e)$ is finite.

The optimized formation problem for the multi-agent system (8) is to find the admissible control policies u_i, $i = 1, \ldots, n$, such that the infinite horizon value function (14) can be minimized.

The control objective. Based on the RL algorithm of the identifier–actor–critic architecture, design the optimized formation control u_i, $i = 1, \ldots, n$, for multi-agent system (8) such that 1) all signals are semiglobally uniformly ultimately bounded (SGUUB); and 2) the leader–follower formation control can be achieved.

Based on the infinite horizon value function (14), the following Hamiltonian function is derived:

$$H\left(e, u, \frac{\partial V}{\partial e}\right) = r(e, u) + \frac{\partial V(e)}{\partial e^T}\dot{e}(t)$$

$$= e^T e + u^T(C \otimes I_m)u + \sum_{i=1}^n \left(\frac{\partial V_i(e_i)}{\partial e_i^T}\dot{e}_i(t)\right)$$

$$= \sum_{i=1}^n \left(\|e_i(t)\|^2 + c_i\|u_i\|^2 + \frac{\partial V_i(e_i)}{\partial e_i^T}\dot{e}_i(t)\right) \quad (16)$$

where $\frac{\partial V(e)}{\partial e}$ and $\frac{\partial V_i(e_i)}{\partial e_i}$ denote the gradient of $V(e(t))$ and $V_i(e_i)$ corresponding to $e(t)$ and $e_i(t)$, respectively.

Let $u^* = [u_1^{*T}, \ldots, u_n^{*T}]^T$ be the optimal formation control, then the optimal value function can be yielded as

$$V^*(e) = \min_{u_{i=1,\cdots,n} \in \Psi(\Omega)} \int_t^\infty r(e, u) d\tau = \int_t^\infty r(e, u^*) d\tau$$

$$= \sum_{i=1}^n V_i^*(e_i) = \sum_{i=1}^n \min_{u_i \in \Psi(\Omega)} \int_t^\infty r_i(e_i, u_i) d\tau$$

$$= \sum_{i=1}^n \int_t^\infty r_i(e_i, u_i^*) d\tau \quad (17)$$

where $V_i^*(e_i) = \int_t^\infty r_i(e_i, u_i^*) d\tau$, $\Omega \subset R^m$ is a compact set containing origin.

Integrating both (16) and (17), the HJB equation is yielded as

$$H\left(e, u^*, \frac{\partial V^*}{\partial e}\right) = r(e, u^*) + \frac{\partial V^*(e)}{\partial e^T}\dot{e}(t)$$

$$= \sum_{i=1}^n \left(\|e_i\|^2 + c_i\|u_i^*\|^2 + \frac{\partial V_i^*(e_i)}{\partial e_i^T}\dot{e}_i(t)\right) = 0. \quad (18)$$

Associated with (13) and (18), the distributed HJB equation can be derived as

$$H_i\left(e_i, u_i^*, \frac{\partial V_i^*}{\partial e_i}\right) = \|e_i\|^2 + c_i\|u_i^*\|^2 + \frac{\partial V_i^*(e_i)}{\partial e_i^T}\left(c_i f_i(x_i) \right.$$

$$\left. + c_i u_i^* - b_i \dot{x}_d(t) - \sum_{j \in \Lambda_i} a_{ij}\dot{x}_j(t)\right) = 0, \quad i = 1, \ldots, n. \quad (19)$$

Obviously, if the distributed HJB equations (19) are held, the HJB equation (18) is held. Assuming the solution of (19) is existent and unique, the following optimal formation control u_i^* can be obtained by solving $\partial H_i(e_i, u_i^*, \frac{\partial V_i^*}{\partial e_i})/\partial u_i^* = 0$:

$$u_i^* = -\frac{1}{2}\frac{\partial V_i^*(e_i)}{\partial e_i}, \quad i = 1, \ldots, n. \quad (20)$$

Substituting (20) into (19) yields

$$\|e_i(t)\|^2 + \frac{\partial V_i^*}{\partial e_i^T}\left(c_i f_i(x_i) - b_i \dot{x}_d(t) - \sum_{j \in \Lambda_i} a_{ij}\dot{x}_j(t)\right)$$

$$- \frac{c_i}{4}\frac{\partial V_i^*}{\partial e_i^T}\frac{\partial V_i^*}{\partial e_i} = 0, \quad i = 1, \ldots, n. \quad (21)$$

In order to achieve the optimal formation control (20), the term $\frac{\partial V_i^*(e_i)}{\partial e_i}$ is required, which is expected to obtain by solving (21). However, due to the unknown dynamics and inherent nonlinearities, the equation is impossible or very difficult to be solved. Therefore, the RL algorithm of the identifier–actor–critic architecture can be considered to realize the control.

B. FLS Identifier Design

Since these dynamic functions $f_i(x_i)$, $i = 1, \ldots, n$, of multi-agent system (8) are unknown, the FLS-based identifiers are established to estimate the unknown functions for achieving the optimized formation scheme.

For $x_i \in \Omega$ where $i = 1, \ldots, n$, the function $f_i(x_i)$ can be approximated by the FLS in the following:

$$f_i(x_i) = \Theta_{fi}^{*T}\varphi_{fi}(x_i) + \varepsilon_{fi}(x_i), \quad i = 1, \ldots, n \quad (22)$$

where $\Theta_{fi}^* \in R^{p_1 \times m}$ is the optimal parameter matrix; $\varphi_{fi}(x_i) \in R^{p_1}$ is the fuzzy basis function vector; p_1 is the fuzzy rule number; $\varepsilon_{fi}(x_i) \in R^m$ is the approximation error satisfying $\|\varepsilon_{fi}(x_i)\| \leq \delta_{fi}$, and δ_{fi} is a positive constant.

Since the optimal parameter matrix Θ_{fi}^* is the unknown constant matrix that cannot be applied directly, it needs to be estimated. Let $\hat{\Theta}_{fi}^T(t)$ denote the estimation, the adaptive identifier

is built as

$$\dot{\hat{x}}_i(t) = -k_i \tilde{x}_i(t) + \hat{\Theta}_{fi}^T(t)\varphi_{fi}(x_i) + u_i,$$
$$i = 1, \dots, n \tag{23}$$

where $\hat{x}_i(t) \in R^m$ is the identifier state, and $\tilde{x}_i(t) = \hat{x}_i(t) - x_i(t)$ is the identification error.

Design the updating law for $\hat{\Theta}_{fi}(t)$ as

$$\dot{\hat{\Theta}}_{fi}(t) = \Gamma_i \left(-\varphi_{fi}(x_i)\tilde{x}_i^T(t) - \sigma_i \hat{\Theta}_{fi}(t) \right),$$
$$i = 1, \dots, n \tag{24}$$

where $\Gamma_i \in R^{p_1 \times p_1}$ is the positive definite gain matrix and σ_i is the positive design parameter.

Based on (8), (22), and (23), the identifier error dynamics can be yielded as

$$\dot{\tilde{x}}_i(t) = -k_i \tilde{x}_i(t) + \tilde{\Theta}_{fi}^T(t)\varphi_{fi}(x_i) - \varepsilon_{fi}(x_i),$$
$$i = 1, \dots, n \tag{25}$$

where $\tilde{\Theta}_{fi}(t) = \hat{\Theta}_{fi}(t) - \Theta_{fi}^*$ is the estimation error.

Theorem 1: If the proposed identifier (23) with updating law (24) is used for identifying the multi-agent (8), then 1) the errors $\tilde{\Theta}_{fi}(t)$ and $\tilde{x}_i(t)$ are SGUUB; 2) the identification error $\tilde{x}_i(t)$ can arrive to the desired accuracy by making the design parameters k_i, $i = 1, \dots, n$, large enough.

Proof: 1) Consider the Lyapunov candidate as following:

$$E_1(t) = \frac{1}{2}\sum_{i=1}^{n}\tilde{x}_i^T(t)\tilde{x}_i(t) + \frac{1}{2}\sum_{i=1}^{n}\text{Tr}\left(\tilde{\Theta}_{fi}^T \Gamma_i^{-1}\tilde{\Theta}_{fi}\right). \tag{26}$$

Taking the time derivative along (24) and (25) is

$$\dot{E}_1(t) = \sum_{i=1}^{n}\tilde{x}_i^T(t)\left(-k_i\tilde{x}_i(t) + \tilde{\Theta}_{fi}^T(t)\varphi_{fi}(x_i) - \varepsilon_{fi}(x_i)\right)$$
$$- \sum_{i=1}^{n}\text{Tr}\left(\tilde{\Theta}_{fi}^T(t)\varphi_{fi}(x_i)\tilde{x}_i^T(t) + \sigma_i\tilde{\Theta}_{fi}^T(t)\hat{\Theta}_{fi}(t)\right). \tag{27}$$

According to the property of trace operator $\text{Tr}(ba^T) = a^T b$ where $a, b \in R^n$, there is the following fact:

$$\text{Tr}\left[\tilde{\Theta}_{fi}^T(t)\varphi_{fi}(x_i)\tilde{x}_i^T(t)\right] = \tilde{x}_i^T(t)\left(\tilde{\Theta}_{fi}^T(t)\varphi_{fi}(x_i)\right). \tag{28}$$

Substituting (28) into (27), we obtain

$$\dot{E}_1(t) = -\sum_{i=1}^{n}k_i\|\tilde{x}_i(t)\|^2 - \sum_{i=1}^{n}\tilde{x}_i^T(t)\varepsilon_{fi}(x_i)$$
$$- \sum_{i=1}^{n}\sigma_i\text{Tr}\left(\tilde{\Theta}_{fi}^T(t)\hat{\Theta}_{fi}(t)\right). \tag{29}$$

According to the Cauchy–Buniakowsky–Schwarz inequality [33] $\left(\sum_{k=1}^{n}a_k b_k\right)^2 \leq \left(\sum_{k=1}^{n}a_k^2\right)\left(\sum_{k=1}^{n}b_k^2\right)$ and Young's inequality [34] $ab \leq \frac{a^2}{2} + \frac{b^2}{2}$, there is the following result:

$$-\tilde{x}_i^T(t)\varepsilon_{fi}(x_i) \leq \frac{1}{2}\|\tilde{x}_i^T(t)\|^2 + \frac{1}{2}\delta_{fi}^2. \tag{30}$$

Based on the fact that $\text{Tr}(\tilde{\Theta}_{fi}^T\hat{\Theta}_{fi}) = \frac{1}{2}\text{Tr}(\tilde{\Theta}_{fi}^T\tilde{\Theta}_{fi}) + \frac{1}{2}\text{Tr}(\hat{\Theta}_{fi}^T\hat{\Theta}_{fi}) - \frac{1}{2}\text{Tr}(\Theta_{fi}^{*T}\Theta_{fi}^*)$, the following equation can be obtained:

$$-\sigma_i\text{Tr}\left(\tilde{\Theta}_{fi}^T(t)\hat{\Theta}_{fi}(t)\right) \leq -\frac{\sigma_i}{2}\text{Tr}\left(\tilde{\Theta}_{fi}^T(t)\tilde{\Theta}_{fi}(t)\right)$$
$$+ \frac{\sigma_i}{2}\text{Tr}\left(\Theta_{fi}^{*T}\Theta_{fi}^*\right). \tag{31}$$

Substituting (30) and (31) into (29) yields

$$\dot{E}_1(t) \leq -\sum_{i=1}^{n}\left(k_i - \frac{1}{2}\right)\|\tilde{x}_i\|^2 - \sum_{i=1}^{n}\frac{\sigma_i}{2}\text{Tr}\left(\tilde{\Theta}_{fi}^T\tilde{\Theta}_{fi}\right) + \beta_1$$
$$\leq -\sum_{i=1}^{n}\left(k_i - \frac{1}{2}\right)\|\tilde{x}_i(t)\|^2 - \sum_{i=1}^{n}\frac{\sigma_i}{2\lambda_{\max}(\Gamma_i^{-1})}$$
$$\times \text{Tr}\left(\tilde{\Theta}_{fi}^T(t)\Gamma_i^{-1}\tilde{\Theta}_{fi}(t)\right) + \beta_1 \tag{32}$$

where $\beta_1 = \frac{1}{2}\sum_{i=1}^{n}(\sigma_i\text{Tr}(\Theta_{fi}^{*T}\Theta_{fi}^*) + \delta_{fi}^2)$; and $\lambda_{\max}(\Gamma_i^{-1})$ denotes the maximal eigenvalue of Γ_i^{-1}.

Let $\alpha_1 = \min\{2(k_1 - \frac{1}{2}), \dots, 2(k_n - \frac{1}{2}), \frac{\sigma_1}{\lambda_{\max}(\Gamma_1^{-1})}, \dots, \frac{\sigma_n}{\lambda_{\max}(\Gamma_n^{-1})}\}$, (32) can be rewritten as

$$\dot{E}_1(t) \leq -\alpha_1 E_1(t) + \beta_1. \tag{33}$$

According to Lemma 4, the following inequality can be obtained:

$$E_1(t) \leq e^{-\alpha_e t}E_1(0) + \frac{\beta_e}{\alpha_e}\left(1 - e^{-\alpha_e t}\right) \tag{34}$$

it implies that the identifier and estimation errors are SGUUB.

2) Let $E_x(t) = \frac{1}{2}\sum_{i=1}^{n}\tilde{x}_i^T(t)\tilde{x}_i(t)$, its time derivative along (25) is

$$\dot{E}_x(t) \leq \sum_{i=1}^{n}\left(-k_i\|\tilde{x}_i\|^2 + \tilde{x}_i^T\tilde{\Theta}_{fi}^T\varphi_{fi}(x_i) - \tilde{x}_i^T\varepsilon_{fi}\right). \tag{35}$$

Inserting the following facts:

$$\tilde{x}_i^T(t)\tilde{\Theta}_{fi}^T(t)\varphi_{fi}(x_i) \leq \frac{1}{2}\|\tilde{x}_i(t)\|^2 + \frac{1}{2}\left\|\tilde{\Theta}_{fi}^T(t)\varphi_{fi}(x_i)\right\|^2,$$
$$-\tilde{x}_i^T(t)\varepsilon_{fi}(x_i) \leq \frac{1}{2}\|\tilde{x}_i(t)\|^2 + \frac{1}{2}\delta_{fi}^2$$

to (35) yields

$$\dot{E}_x(t) \leq -\sum_{i=1}^{n}(k_i - 1)\|\tilde{x}_i(t)\|^2 + \psi_x(t) \tag{36}$$

where $\psi_x(t) = \frac{1}{2}\sum_{i=1}^{n}(\|\tilde{\Theta}_{fi}^T(t)\varphi_{fi}(x_i)\|^2 + \delta_{fi}^2)$.

Since these estimation errors $\tilde{\Theta}_{f1}^T(t), \dots, \tilde{\Theta}_{fn}^T(t)$ are bounded, which are proven by part 1, the term $\psi_x(t)$ is bounded. Let $\alpha_2 = \min_{i=1,\dots,n}\{k_i - 1\}$ and $\beta_2 = \sup_{t \geq 0}\{\psi_x(t)\}$, (36) becomes

$$\dot{E}_x(t) \leq -\alpha_2 E_x(t) + \beta_2. \tag{37}$$

Applying Lemma (4), we obtain the following equation:

$$E_x(t) \leq e^{-\alpha_2 t}E_x(0) + \frac{\beta_2}{\alpha_2}\left(1 - e^{-\beta_2 t}\right). \tag{38}$$

The above-mentioned inequality means that the identifier error can arrive the desired accuracy by making α_2 large enough. $\square$

C. Optimized Formation Control Design

Since the multi-agent dynamic function $f_i(x_i)$ is unknown, the identifier (23) plays an essential role in the formation control design. Define the identifier tracking and identifier formation errors as

$$\hat{z}_i(t) = \hat{x}_i(t) - x_d(t) - \eta_i,$$

$$\hat{e}_i(t) = \sum_{j \in \Lambda_i} a_{ij} \left(\hat{x}_i(t) - \eta_i - \hat{x}_j + \eta_j \right) + b_i \hat{z}_i(t). \quad (39)$$

Based on the identifier dynamic (23), the following error dynamics can be yielded:

$$\dot{\hat{z}}_i(t) = -k_i \tilde{x}_i(t) + \hat{\Theta}_{fi}^T(t)\varphi_{fi}(x_i) - \dot{x}_d(t) + u_i, \quad (40)$$

$$\dot{\hat{e}}_i(t) = -k_i c_i \tilde{x}_i(t) + c_i \hat{\Theta}_{fi}^T(t)\varphi_{fi}(x_i) + c_i u_i - b_i \dot{x}_d$$
$$- \sum_{j \in \Lambda_i} a_{ij} \dot{\hat{x}}_j(t), \quad i = 1, \dots, n. \quad (41)$$

Similar to (14)–(19), the optimal value function for the error dynamic (41) is

$$V^*(\hat{e}) = \min_{u_{i=1,\dots,n} \in \Psi(\Omega)} \int_t^\infty r(\hat{e}(\tau), u(\hat{e})) \, d\tau$$

$$= \sum_{i=1}^n V_i^*(\hat{e}_i) = \sum_{i=1}^n \min_{u_i \in \Psi(\Omega)} \int_t^\infty r_i(\hat{e}_i(\tau), u_i(\hat{e}_i)) \, d\tau$$

$$= \sum_{i=1}^n \int_t^\infty r_i(\hat{e}_i(\tau), u_i^*(\hat{e}_i)) \, d\tau \quad (42)$$

where $\hat{e}(t) = [\hat{e}_1^T(t), \hat{e}_2^T(t), \dots, \hat{e}_n^T(t)]^T$. Then the distributed HJB equation associated with (41) can be yielded as

$$H_i\left(\hat{e}_i, u_i^*, \frac{\partial V_i^*}{\partial \hat{e}_i}\right) = \|\hat{e}_i(t)\|^2 + c_i \|u_i^*\|^2 + \frac{\partial V_i^*(\hat{e}_i)}{\partial \hat{e}_i^T} \dot{\hat{e}}_i$$

$$= \|\hat{e}_i(t)\|^2 + c_i \|u_i^*\|^2 + \frac{\partial V_i^*(\hat{e}_i)}{\partial \hat{e}_i^T}\left(-k_i c_i \tilde{x}_i(t) + c_i u_i^*\right.$$

$$\left. + c_i \hat{\Theta}_{fi}^T(t)\varphi_{fi}(x_i) - b_i \dot{x}_d - \sum_{j \in \Lambda_i} a_{ij} \dot{\hat{x}}_j(t)\right) = 0,$$

$$i = 1, \dots, n. \quad (43)$$

Assume the solution of (43) to be existent and unique. By solving $\partial H_i(\hat{e}_i, u_i^*, \frac{\partial V_i^*}{\partial \hat{e}_i})/\partial u_i^* = 0$, the optimal formation control u_i^* can be obtained as

$$u_i^* = -\frac{1}{2} \frac{\partial V_i^*(\hat{e}_i)}{\partial \hat{e}_i}, \quad i = 1, \dots, n. \quad (44)$$

Segment the optimal value function (42) into two parts as

$$V_i^*(\hat{e}_i) = \gamma_i \|\hat{e}_i(t)\|^2 + V_i^o(\hat{e}_i), \quad i = 1, \dots, n \quad (45)$$

where γ_i is a positive design constant, and $V_i^o(\hat{e}_i) = -\gamma_i \|\hat{e}_i(t)\|^2 + V_i^*(\hat{e}_i)$. Inserting (45) into (44), the optimal formation control can become

$$u_i^* = -\gamma_i \hat{e}_i(t) - \frac{1}{2} \frac{\partial V_i^o}{\partial \hat{e}_i}, \quad i = 1, \dots, n. \quad (46)$$

Since $V_i^o(\hat{e}_i)$ is the continuous function, for $\hat{e}_i \in \Omega$ where $i = 1, \dots, n$, $V_i^o(\hat{e}_i)$ can be approximated by FLS as

$$V_i^o(\hat{e}_i) = \Theta_i^{*T} \varphi_i(\hat{e}_i) + \varepsilon_i(\hat{e}_i), \quad i = 1, \dots, n \quad (47)$$

where $\Theta_i^* \in R^{p_2}$ is the optimal parameter matrix; $\varphi_i(\hat{e}_i) \in R^{p_2}$ is the fuzzy basis function vector; p_2 is the fuzzy rule number; and $\varepsilon_i(\hat{e}_i) \in R$ is the approximation error to satisfy $|\varepsilon_i(\hat{e}_i)| \leq \delta_i$ where δ_i is a constant.

Based on the FLS approximation (47), the optimal value function (45) and optimal control (46) can be rewritten as

$$V_i^*(\hat{e}_i) = \gamma_i \|\hat{e}_i(t)\|^2 + \Theta_i^{*T} \varphi_i(\hat{e}_i) + \varepsilon_i(\hat{e}_i), \quad (48)$$

$$u_i^* = -\gamma_i \hat{e}_i(t) - \frac{1}{2} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \Theta_i^* - \frac{1}{2} \frac{\partial \varepsilon_i(\hat{e}_i)}{\partial \hat{e}_i},$$

$$i = 1, \dots, n \quad (49)$$

where $\frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i}$ and $\frac{\partial \varepsilon_i(\hat{e}_i)}{\partial \hat{e}_i}$ are the gradients with respect to $\hat{e}_i$.

Substituting (48) and (49) into (43), we obtain the following equation:

$$H_i\left(\hat{e}_i, u_i^*, \frac{\partial V_i^*}{\partial \hat{e}_i}\right) = -(\gamma_i^2 c_i - 1)\|\hat{e}_i(t)\|^2 + 2\gamma_i \hat{e}_i^T(t)$$

$$\times \left(c_i \hat{\Theta}_{fi}^T \varphi_{fi}(x_i) - k_i c_i \tilde{x}_i(t) - b_i \dot{x}_d - \sum_{j \in \Lambda_i} a_{ij} \dot{\hat{x}}_j(t)\right)$$

$$+ \Theta_i^{*T} \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T}\left(c_i \hat{\Theta}_{fi}^T \varphi_{fi}(x_i) - \gamma_i c_i \hat{e}_i(t) - k_i c_i \tilde{x}_i(t)\right.$$

$$\left. - b_i \dot{x}_d(t) - \sum_{j \in \Lambda_i} a_{ij} \dot{\hat{x}}_j(t)\right) - \frac{c_i}{4}\left\|\frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \Theta_i^*\right\|^2$$

$$+ \epsilon_i(t) = 0 \quad (50)$$

where

$$\epsilon_i(t) = \frac{\partial \varepsilon_i(\hat{e}_i)}{\partial \hat{e}_i^T}\left(c_i u_i^* - k_i c_i \tilde{x}_i(t) + c_i \hat{\Theta}_{fi}^T \varphi_{fi}(x_i) - b_i \dot{x}_d\right.$$

$$\left. - \sum_{j \in \Lambda_i} a_{ij} \dot{\hat{x}}_j(t)\right) + \frac{c_i}{4}\left\|\frac{\partial \varepsilon_i(\hat{e}_i)}{\partial \hat{e}_i}\right\|^2.$$

The term $\epsilon_i(t)$ is bounded because all terms are bounded.

Since the optimal parameter matrix Θ_i^* is unknown, the optimal formation controller (49) cannot be applied directly. In order to obtain the available control scheme, the following actor–critic RL algorithm is constructed based on the FLS approximation (47), of which actor and critic FLSs are utilized to implement the control behavior and evaluate the control performance, respectively:

$$\hat{V}_i^*(\hat{e}_i) = \gamma_i \|\hat{e}_i(t)\|^2 + \hat{\Theta}_{ci}^T(t)\varphi_i(\hat{e}_i), \quad (51)$$

$$u_i = -\gamma_i \hat{e}_i(t) - \frac{1}{2} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \hat{\Theta}_{ai}(t), \quad i = 1, \dots, n \quad (52)$$

where $\hat{V}_i^*(\hat{e}_i)$ denotes the estimations of $V_i^*(\hat{e}_i)$; and $\hat{\Theta}_{ci}(t) \in R^{p_2}$ and $\hat{\Theta}_{ai}(t) \in R^{p_2}$ are the critic and actor parameter vectors, respectively.

Using (51) and (52), the approximated HJB equation can be obtained as

$$
H_i\left(\hat{e}_i, u_i, \frac{\partial \hat{V}_i^*}{\partial \hat{e}_i}\right) = \|\hat{e}_i\|^2 + c_i \left\| -\gamma_i \hat{e}_i - \frac{1}{2}\frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i}\hat{\Theta}_{ai}(t)\right\|^2
$$
$$
+ \left(2\gamma_i \hat{e}_i^T + \hat{\Theta}_{ci}^T(t)\frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T}\right)\left(c_i \hat{\Theta}_{fi}^T(t)\varphi_{fi}(x_i) - k_i c_i \tilde{x}_i(t)\right.
$$
$$
\left. -\gamma_i c_i \hat{e}_i - \frac{c_i}{2}\frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i}\hat{\Theta}_{ai}^T(t) - b_i \dot{x}_d - \sum_{j\in\Lambda_i} a_{ij}\dot{\hat{x}}_j\right),
$$
$$
i = 1, \ldots, n. \tag{53}
$$

Define the Bellman residual error $\phi_i(t)$ as

$$
\phi_i(t) = H_i\left(\hat{e}_i, u_i, \frac{\partial \hat{V}_i^*}{\partial \hat{e}_i}\right) - H_i\left(\hat{e}_i, u_i^*, \frac{\partial V_i^*}{\partial \hat{e}_i}\right)
$$
$$
= H_i\left(\hat{e}_i, u_i, \frac{\partial \hat{V}_i^*}{\partial \hat{e}_i}\right), \quad i = 1, \ldots, n. \tag{54}
$$

Let $\Phi_i(t) = \frac{1}{2}\phi_i^2(t)$, the critic updating law can be yielded based on the gradient descent algorithm for minimizing the Bellman residual error:

$$
\dot{\hat{\Theta}}_{ci}(t) = -\frac{\kappa_{ci}}{1 + \|\xi_i(t)\|^2}\frac{\partial \Phi_i(t)}{\partial \hat{\Theta}_{ci}(t)}
$$
$$
= -\frac{\kappa_{ci}\xi_i(t)}{1 + \|\xi_i(t)\|^2}\left(\xi_i^T(t)\hat{\Theta}_{ci}(t) - (\gamma_i^2 c_i - 1)\|\hat{e}_i(t)\|^2\right.
$$
$$
+ 2\gamma_i \hat{e}_i^T\left(c_i \hat{\Theta}_{fi}^T(t)\varphi_{fi}(x_i) - k_i c_i \tilde{x}_i - b_i \dot{x}_d - \sum_{j\in\Lambda_i} a_{ij}\dot{\hat{x}}_j\right)
$$
$$
\left. + \frac{c_i}{4}\left\|\frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i}\hat{\Theta}_{ai}(t)\right\|^2\right), \quad i = 1, \ldots, n \tag{55}
$$

where $\kappa_{ci} > 0$ is the critic learning rate; and

$$
\xi_i(t) = \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T}\left(c_i \hat{\Theta}_{fi}^T(t)\varphi_{fi}(x_i) - k_i c_i \tilde{x}_i - \gamma_i c_i \hat{e}_i\right.
$$
$$
\left. -\frac{c_i}{2}\frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i}\hat{\Theta}_{ai}(t) - b_i \dot{x}_d(t) - \sum_{j\in\Lambda_i} a_{ij}\dot{\hat{x}}_j(t)\right).
$$

The actor weight updating law is designed as

$$
\dot{\hat{\Theta}}_{ai}(t) = \frac{1}{2}\frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T}\hat{e}_i(t) - \kappa_{ai}c_i \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T}\frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i}\hat{\Theta}_{ai}(t)
$$
$$
+ \frac{\kappa_{ci}c_i}{4\left(1 + \|\xi_i(t)\|^2\right)}\frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T}\frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i}
$$
$$
\times \hat{\Theta}_{ai}(t)\xi_i^T(t)\hat{\Theta}_{ci}(t), i = 1, \ldots, n \tag{56}
$$

where $\kappa_{ai} > 0$ is the actor learning rate.

Assumption 1: [28] Persistence of excitation (PE): the signs of $\xi_i(t)\xi_i^T(t)$, $i = 1, 2, \ldots, n$, are required persistent excitation over the interval $[t, t + \bar{t}_i]$, i.e., there exist constants $\varsigma_i > 0$, $\zeta_i > 0$, $\bar{t}_i > 0$ for all t satisfying the following condition:

$$
\varsigma_i I_{p_2} \leq \xi_i(t)\xi_i^T(t) \leq \zeta_i I_{p_2} \tag{57}
$$

where $I_{p_2} \in R^{p_2 \times p_2}$ is the identity matrix.

D. Stability Analysis

Theorem 2: Consider the multi-agent system (8) with bounded initial conditions and reference signal. If the optimized multi-agent formation control (52) is performed based on the identifier–critic–actor RL algorithm, where the identifier, actor, and critic are online trained by the adaptive laws (24), (55), and (56), respectively, then by choosing appropriate design parameters, the optimized formation control can guarantee that

1) all error signals are SGUUB; and
2) the leader–follower formation control can be achieved.

Proof: 1) Choose the Lyapunov function candidate as

$$
E(t) = \frac{1}{2}\hat{z}^T(t)(\tilde{L}\otimes I_m)\hat{z}(t) + \frac{1}{2}\sum_{i=1}^n \tilde{\Theta}_{ai}^T(t)\tilde{\Theta}_{ai}(t)
$$
$$
+ \frac{1}{2}\sum_{i=1}^n \tilde{\Theta}_{ci}^T(t)\tilde{\Theta}_{ci}(t) \tag{58}
$$

where $\tilde{\Theta}_{ai}(t) = \hat{\Theta}_{ai}(t) - \Theta^*$, $\tilde{\Theta}_{ci}(t) = \hat{\Theta}_{ci}(t) - \Theta^*$. The time derivative along (40), (55), and (56) is

$$
\dot{E}(t) = \sum_{i=1}^n \hat{e}_i^T(t)\left(-k_i \tilde{x}_i(t) + \hat{\Theta}_{fi}^T(t)\varphi_{fi}(x_i) - \dot{x}_d(t) + u_i\right)
$$
$$
+ \sum_{i=1}^n \tilde{\Theta}_{ai}^T(t)\left(\frac{1}{2}\frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T}\hat{e}_i - \kappa_{ai}c_i \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T}\frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i}\hat{\Theta}_{ai}\right.
$$
$$
\left. + \frac{\kappa_{ci}c_i}{4\left(1 + \|\xi_i(t)\|^2\right)}\frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T}\frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i}\hat{\Theta}_{ai}(t)\xi_i^T(t)\hat{\Theta}_{ci}(t)\right)
$$
$$
+ \sum_{i=1}^n \tilde{\Theta}_{ci}^T(t)\left(-\frac{\kappa_{ci}\xi_i(t)}{1 + \|\xi_i\|^2}\left(\xi_i^T(t)\hat{\Theta}_{ci} - (\gamma_i^2 c_i - 1)\|\hat{e}_i\|^2\right.\right.
$$
$$
+ 2\gamma_i \hat{e}_i^T(t)\left(c_i \hat{\Theta}_{fi}^T(t)\varphi_{fi}(x_i) - k_i c_i \tilde{x}_i - b_i \dot{x}_d - \sum_{j\in\Lambda_i} a_{ij}\dot{\hat{x}}_j\right)
$$
$$
\left.\left. + \frac{c_i}{4}\left\|\frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i}\hat{\Theta}_{ai}(t)\right\|^2\right)\right). \tag{59}
$$

According to Young's and Cauchy–Buniakowsky–Schwarz inequalities, there are the following facts:

$$- k_i \hat{e}_i^T(t) \tilde{x}_i(t) \leq k_i \|\hat{e}_i(t)\|^2 + \frac{k_i}{4} \|\tilde{x}_i(t)\|^2,$$

$$\hat{e}_i^T(t) \hat{\Theta}_{fi}^T(t) \varphi_{fi}(x_i) \leq \frac{1}{2} \|\hat{e}_i(t)\|^2 + \frac{1}{2} \left\| \hat{\Theta}_{fi}^T(t) \varphi_{fi} \right\|^2,$$

$$- \hat{e}_i^T(t) \dot{x}_d(t) \leq \frac{1}{2} \|\hat{e}_i(t)\|^2 + \frac{1}{2} \|\dot{x}_d(t)\|^2. \tag{60}$$

Inserting the above-mentioned inequalities and control law (52) into (59), we obtain

$$\dot{E}(t) \leq - \sum_{i=1}^{n} (\gamma_i - k_i - 1) \|\hat{e}_i\|^2 + \sum_{i=1}^{n} \left(-\frac{1}{2} \hat{e}_i^T \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \hat{\Theta}_{ai} \right.$$

$$+ \frac{1}{2} \tilde{\Theta}_{ai}^T \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \hat{e}_i - \kappa_{ai} c_i \tilde{\Theta}_{ai}^T(t) \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \hat{\Theta}_{ai}(t)$$

$$\left. + \frac{\kappa_{ci} c_i}{4 \left(1 + \|\xi_i\|^2 \right)} \tilde{\Theta}_{ai}^T(t) \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \hat{\Theta}_{ai}(t) \xi_i^T \hat{\Theta}_{ci}(t) \right)$$

$$+ \sum_{i=1}^{n} \tilde{\Theta}_{ci}^T(t) \left(-\frac{\kappa_{ci} \xi_i}{1 + \|\xi_i\|^2} \left(\xi_i^T \hat{\Theta}_{ci}(t) - (\gamma_i^2 c_i - 1) \|\hat{e}_i\|^2 \right. \right.$$

$$+ 2\gamma_i \hat{e}_i^T \left(\hat{\Theta}_{fi}^T(t) \varphi_{fi}(x_i) - k_i c_i \tilde{x}_i - b_i \dot{x}_d - \sum_{j \in \Lambda_i} a_{ij} \dot{\hat{x}}_j \right)$$

$$\left. + \frac{c_i}{4} \left\| \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \hat{\Theta}_{ai}(t) \right\|^2 \right) + \sum_{i=1}^{n} \left(\frac{k_i}{4} \|\tilde{x}_i\|^2 + \frac{1}{2} \|\dot{x}_d\|^2 \right.$$

$$\left. + \frac{1}{2} \left\| \hat{\Theta}_{fi}^T(t) \varphi_{fi}(x_i) \right\|^2 \right). \tag{61}$$

Based on the fact that $\tilde{\Theta}_{ai}(t) = \hat{\Theta}_{ai}(t) - \Theta_i^*$, there are the following equations:

$$\tilde{\Theta}_{ai}^T(t) \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \hat{e}_i - \hat{e}_i^T(t) \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \hat{\Theta}_{ci} = -\hat{e}_i^T \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \Theta_i^*,$$

$$- \kappa_{ai} c_i \tilde{\Theta}_{ai}^T(t) \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \hat{\Theta}_{ai}(t) = -\frac{\kappa_{ai} c_i}{2} \tilde{\Theta}_{ai}^T(t)$$

$$\frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \tilde{\Theta}_{ai}(t) - \frac{\kappa_{ai} c_i}{2} \hat{\Theta}_{ai}^T(t) \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i}$$

$$\hat{\Theta}_{ai}(t) + \frac{\kappa_{ai} c_i}{2} \Theta_i^{*T} \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \Theta_i^*.$$

Substituting the above-mentioned equations into (61) yields

$$\dot{E}(t) \leq - \sum_{i=1}^{n} (\gamma_i - k_i - 1) \|\hat{e}_i\|^2 - \frac{1}{2} \sum_{i=1}^{n} \hat{e}_i^T \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \Theta_i^*$$

$$- \sum_{i=1}^{n} \frac{\kappa_{ai} c_i}{2} \tilde{\Theta}_{ai}^T \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \tilde{\Theta}_{ai} - \sum_{i=1}^{n} \frac{\kappa_{ai} c_i}{2} \hat{\Theta}_{ai}^T(t)$$

$$\frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \hat{\Theta}_{ai}(t) + \sum_{i=1}^{n} \frac{\kappa_{ci} c_i}{4 \left(1 + \|\xi_i\|^2 \right)}$$

$$\tilde{\Theta}_{ai}^T(t) \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \hat{\Theta}_{ai}(t) \xi_i^T(t) \hat{\Theta}_{ci}(t) + \sum_{i=1}^{n} \tilde{\Theta}_{ci}^T(t)$$

$$\left(-\frac{\kappa_{ci} \xi_i(t)}{1 + \|\xi_i(t)\|^2} \left(\xi_i^T \hat{\Theta}_{ci}(t) - (\gamma_i^2 c_i - 1) \|\hat{e}_i\|^2 + 2\gamma_i \hat{e}_i^T(t) \right. \right.$$

$$\left. \left(\hat{\Theta}_{fi}^T(t) \varphi_{fi}(x_i) - k_i c_i \tilde{x}_i - b_i \dot{x}_d - \sum_{j \in \Lambda_i} a_{ij} \dot{\hat{x}}_j \right) \right.$$

$$\left. + \frac{c_i}{4} \left\| \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \hat{\Theta}_{ai}(t) \right\|^2 \right) + \sum_{i=1}^{n} \left(\frac{k_i}{4} \|\tilde{x}_i(t)\|^2 + \frac{1}{2} \|\dot{x}_d(t)\|^2 \right.$$

$$\left. + \frac{1}{2} \left\| \hat{\Theta}_{fi}^T(t) \varphi_{fi}(x_i) \right\|^2 + \frac{\kappa_{ai} c_i}{2} \Theta_i^{*T} \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \Theta_i^* \right). \tag{62}$$

According to (50), the following equation can be obtained:

$$- (\gamma_i^2 c_i - 1) \|\hat{e}_i(t)\|^2 + 2\gamma_i \hat{e}_i^T(t) \left(c_i \hat{\Theta}_{fi}^T(t) \varphi_{fi}(x_i) - k_i c_i \tilde{x}_i \right.$$

$$\left. - b_i \dot{x}_d(t) - \sum_{j \in \Lambda_i} a_{ij} \dot{\hat{x}}_j(t) \right) = -\xi_i^T(t) \Theta_i^* - \frac{c_i}{2} \Theta_i^{*T} \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T}$$

$$\frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \hat{\Theta}_{ai}(t) + \frac{c_i}{4} \left\| \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \Theta_i^* \right\|^2 - \epsilon_i(t). \tag{63}$$

Applying (63) and the fact that

$$- \frac{1}{2} \hat{e}_i^T(t) \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \Theta_i^* \leq \|\hat{e}_i\|^2 + \left\| \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \Theta_i^* \right\|^2 \tag{64}$$

(62) can be rewritten as

$$\dot{E}(t) \leq - \sum_{i=1}^{n} (\gamma_i - k_i - 2) \|\hat{e}_i\|^2 - \sum_{i=1}^{n} \frac{\kappa_{ai} c_i}{2} \tilde{\Theta}_{ai}^T \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T}$$

$$\frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \tilde{\Theta}_{ai} - \sum_{i=1}^{n} \frac{\kappa_{ai} c_i}{2} \hat{\Theta}_{ai}^T \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \hat{\Theta}_{ai}$$

$$+ \sum_{i=1}^{n} \frac{\kappa_{ci} c_i}{4 \left(1 + \|\xi_i\|^2 \right)} \tilde{\Theta}_{ai}^T(t) \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \hat{\Theta}_{ai}(t)$$

$$\xi_i^T(t) \hat{\Theta}_{ci}(t) + \sum_{i=1}^{n} \tilde{\Theta}_{ci}^T(t) \left(-\frac{\kappa_{ci} \xi_i(t)}{1 + \|\xi_i(t)\|^2} \left(\xi_i^T(t) \tilde{\Theta}_{ci}(t) \right. \right.$$

$$+ \frac{c_i}{4} \left\| \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \hat{\Theta}_{ai}(t) \right\|^2 - \frac{c_i}{2} \Theta_i^{*T} \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i}$$

$$\left. \left. \hat{\Theta}_{ai}(t) + \frac{c_i}{4} \left\| \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \Theta_i^* \right\|^2 - \epsilon_i \right) \right) + \sum_{i=1}^{n} \left(\frac{k_i}{4} \|\tilde{x}_i\|^2 \right.$$

$$+ \frac{1}{2} \left\| \hat{\Theta}_{fi}^T(t)\varphi_{fi}(x_i) \right\|^2 + \frac{\kappa_{ai}c_i + 2}{2} \left\| \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \Theta_i^* \right\|^2$$

$$+ \frac{1}{2} \left\| \dot{x}_d \right\|^2 \Bigg). \tag{65}$$

Using the fact that

$$\frac{c_i}{4} \left\| \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \hat{\Theta}_{ai}(t) \right\|^2 - \frac{c_i}{2} \Theta_i^{*T} \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \hat{\Theta}_{ai}(t)$$

$$+ \frac{c_i}{4} \left\| \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \Theta_i^* \right\|^2 = \frac{c_i}{4} \tilde{\Theta}_{ai}^T(t) \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \hat{\Theta}_{ai}(t)$$

$$- \frac{c_i}{4} \Theta_i^{*T} \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \tilde{\Theta}_{ai}(t) \tag{66}$$

(65) can be rewritten as

$$\dot{E}(t) \leq - \sum_{i=1}^n (\gamma_i - k_i - 2) \|\hat{e}_i(t)\|^2 - \sum_{i=1}^n \frac{\kappa_{ai}c_i}{2} \tilde{\Theta}_{ai}^T(t)$$

$$\frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \tilde{\Theta}_{ai}(t) - \sum_{i=1}^n \frac{\kappa_{ci}}{1 + \|\xi_i\|^2} \tilde{\Theta}_{ci}^T(t)\xi_i\xi_i^T \tilde{\Theta}_{ci}(t)$$

$$+ \sum_{i=1}^n \frac{\kappa_{ci}c_i}{4\left(1 + \|\xi_i\|^2\right)} \tilde{\Theta}_{ai}^T(t) \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \hat{\Theta}_{ai}(t)\xi_i^T \hat{\Theta}_{ci}$$

$$- \sum_{i=1}^n \frac{c_i\kappa_{ci}}{4\left(1 + \|\xi_i\|^2\right)} \tilde{\Theta}_{ci}^T(t)\xi_i\tilde{\Theta}_{ai}^T(t) \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \hat{\Theta}_{ai}$$

$$+ \sum_{i=1}^n \frac{c_i\kappa_{ci}}{4\left(1 + \|\xi_i\|^2\right)} \tilde{\Theta}_{ci}^T(t)\xi_i\Theta_i^{*T} \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \tilde{\Theta}_{ai}(t)$$

$$+ \sum_{i=1}^n \frac{\kappa_{ci}}{1 + \|\xi_i\|^2} \tilde{\Theta}_{ci}^T(t)\xi_i\epsilon_i(t) - \sum_{i=1}^n \frac{\kappa_{ai}c_i}{2} \hat{\Theta}_{ai}^T(t) \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T}$$

$$\frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \hat{\Theta}_{ai}(t) + \sum_{i=1}^n \left(\frac{k_i}{4} \|\tilde{x}_i(t)\|^2 + \frac{1}{2} \left\| \hat{\Theta}_{fi}^T \varphi_{fi}(x_i) \right\|^2 \right.$$

$$\left. + \frac{1}{2} \|\dot{x}_d\|^2 + \frac{\kappa_{ai}c_i + 2}{2} \left\| \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \Theta_i^* \right\|^2 \right). \tag{67}$$

Substituting the facts that

$$\sum_{i=1}^n \frac{\kappa_{ci}c_i}{4\left(1 + \|\xi_i\|^2\right)} \tilde{\Theta}_{ai}^T(t) \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \hat{\Theta}_{ai}(t)\xi_i^T \hat{\Theta}_{ci}$$

$$- \sum_{i=1}^n \frac{c_i\kappa_{ci}}{4\left(1 + \|\xi_i\|^2\right)} \tilde{\Theta}_{ci}^T(t)\xi_i\tilde{\Theta}_{ai}^T(t) \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \hat{\Theta}_{ai}(t)$$

$$= \sum_{i=1}^n \frac{c_i\kappa_{ci}}{4\left(1 + \|\xi_i\|^2\right)} \tilde{\Theta}_{ai}^T(t) \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \Theta_i^{*T}\xi_i \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \hat{\Theta}_{ai}(t),$$

$$\frac{\kappa_{ci}}{1 + \|\xi_i\|^2} \tilde{\Theta}_{ci}^T(t)\xi_i\epsilon_i(t) \leq \frac{\kappa_{ci}}{2(1 + \|\xi_i\|^2)} \tilde{\Theta}_{ci}^T(t)\xi_i\xi_i^T \tilde{\Theta}_{ci}(t)$$

$$+ \frac{\kappa_{ci}}{2(1 + \|\xi_i\|^2)} \epsilon_i^2(t)$$

into (67) yields

$$\dot{E}(t) \leq - \sum_{i=1}^n (\gamma_i - k_i - 2) \|\hat{e}_i(t)\|^2 - \sum_{i=1}^n \frac{\kappa_{ai}c_i}{2} \tilde{\Theta}_{ai}^T(t)$$

$$\frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \tilde{\Theta}_{ai}(t) - \sum_{i=1}^n \frac{\kappa_{ci}}{2\left(1 + \|\xi_i\|^2\right)} \tilde{\Theta}_{ci}^T\xi_i\xi_i^T \tilde{\Theta}_{ci}$$

$$+ \sum_{i=1}^n \frac{c_i\kappa_{ci}}{4\left(1 + \|\xi_i\|^2\right)} \tilde{\Theta}_{ai}^T(t) \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \Theta_i^{*T}\xi_i \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \hat{\Theta}_{ai}(t)$$

$$+ \sum_{i=1}^n \frac{c_i\kappa_{ci}}{4\left(1 + \|\xi_i\|^2\right)} \tilde{\Theta}_{ci}^T(t)\xi_i\Theta_i^{*T} \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \tilde{\Theta}_{ai}(t)$$

$$- \sum_{i=1}^n \frac{\kappa_{ai}c_i}{2} \hat{\Theta}_{ai}^T(t) \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \hat{\Theta}_{ai}(t) + \psi_e(t) \tag{68}$$

where

$$\psi_e(t) = \sum_{i=1}^n \left(\frac{k_i}{4} \|\tilde{x}_i(t)\|^2 + \frac{1}{2} \|\dot{x}_d(t)\|^2 \right.$$

$$+ \frac{\kappa_{ci}}{2(1 + \|\xi_i(t)\|^2)} \epsilon_i^2(t) + \frac{1}{2} \left\| \hat{\Theta}_{fi}^T(t)\varphi_{fi} \right\|^2$$

$$\left. + \left(1 + \frac{\kappa_{ai}c_i}{2}\right) \left\| \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \Theta_i^* \right\|^2 \right).$$

Using Young's and Cauchy–Buniakowsky–Schwarz inequalities, we obtain the following results:

$$\frac{\kappa_{ci}c_i}{4\left(1 + \|\xi_i(t)\|^2\right)} \tilde{\Theta}_{ai}^T(t) \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \Theta_i^{*T}\xi_i(t) \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \hat{\Theta}_{ai}(t)$$

$$\leq \frac{c_i}{32} \tilde{\Theta}_{ai}^T(t) \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \Theta_i^{*T}\xi_i(t)\xi_i^T(t)\Theta_i^* \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \tilde{\Theta}_{ai}(t)$$

$$+ \frac{\kappa_{ci}^2 c_i}{2} \hat{\Theta}_{ai}^T(t) \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \hat{\Theta}_{ai}(t),$$

$$\frac{\kappa_{ci}c_i}{4\left(1 + \|\xi_i(t)\|^2\right)} \tilde{\Theta}_{ci}^T(t)\xi_i(t)\Theta_i^{*T} \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \tilde{\Theta}_{ai}(t)$$

$$\leq \frac{c_i}{32\left(1 + \|\xi_i(t)\|^2\right)} \tilde{\Theta}_{ci}^T(t)\xi_i(t)\Theta_i^{*T} \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i}$$

$$\Theta_i^*\xi_i^T(t)\tilde{\Theta}_{ci}(t) + \frac{\kappa_{ci}^2 c_i}{2} \tilde{\Theta}_{ai}^T(t) \frac{\partial \varphi_i(\hat{e}_i)}{\partial \hat{e}_i^T} \frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i} \tilde{\Theta}_{ai}(t).$$

Inserting the above-mentioned facts into (68) yields

$$\dot{E}(t) \leq -\sum_{i=1}^{n}(\gamma_i - k_i - 2)\|\hat{e}_i\|^2 - \sum_{i=1}^{n}\left(\frac{\kappa_{ai}c_i}{2} - \frac{\kappa_{ci}^2 c_i}{2}\right.$$

$$\left. -\frac{c_i}{32}\Theta_i^{*T}\xi_i(t)\xi_i^T(t)\Theta_i^*\right)\tilde{\Theta}_{ai}^T(t)\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\tilde{\Theta}_{ai}(t)$$

$$-\sum_{i=1}^{n}\frac{1}{\left(1+\|\xi_i\|^2\right)}\left(\frac{\kappa_{ci}}{2} - \frac{c_i}{32}\Theta_i^{*T}\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\Theta_i^*\right)$$

$$\tilde{\Theta}_{ci}^T(t)\xi_i\xi_i^T\tilde{\Theta}_{ci}(t) - \sum_{i=1}^{n}\left(\frac{\kappa_{ai}c_i}{2} - \frac{\kappa_{ci}^2 c_i}{2}\right)\hat{\Theta}_{ai}^T(t)\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}$$

$$\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}(t) + \psi_e(t). \tag{69}$$

Make the design parameters to satisfy the following conditions:

$$\gamma_i \geq k_i + 2, \kappa_{ci} \geq \frac{c_i}{16}\Theta_i^{*T}\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\Theta_i^*,$$

$$\kappa_{ai} \geq \kappa_{ci}^2 + \frac{\zeta_i}{16}\Theta_i^{*T}\Theta_i^*. \tag{70}$$

Based on the PE condition (see Assumption 1), (69) can be written as

$$\dot{E}(t) \leq -\sum_{i=1}^{n}(\gamma_i - k_i - 2)\|\hat{e}_i(t)\|^2 - \sum_{i=1}^{n}\left(\frac{\kappa_{ai}c_i}{2} - \frac{\kappa_{ci}^2 c_i}{2}\right.$$

$$\left. -\frac{\zeta_i c_i}{32}\Theta_i^{*T}\Theta_i^*\right)\lambda_i^{\min}\tilde{\Theta}_{ai}^T(t)\tilde{\Theta}_{ai}(t) - \sum_{i=1}^{n}\left(\frac{\kappa_{ci}\varsigma_i}{2} - \frac{\lambda_i^{\max}\varsigma_i c_i}{32}\right.$$

$$\left.\Theta_i^{*T}\Theta_i^*\right)\tilde{\Theta}_{ci}^T(t)\tilde{\Theta}_{ci}(t) + \psi_e(t) \tag{71}$$

where $\lambda_i^{\max}$ and $\lambda_i^{\min}$ are the maximum and minimum eigenvalues of $\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}$.

Let $\gamma = \min\limits_{i=1,\ldots,n}\{\gamma_i - k_i - 2\}$, $\kappa_a = \min\limits_{i=1,\ldots,n}\{(\frac{\kappa_{ai}c_i}{2} - \frac{\kappa_{ci}^2 c_i}{2}$ $-\frac{\zeta_i c_i}{32}\Theta_i^{*T}\Theta_i^*)\lambda_i^{\min}\}$, $\kappa_c = \min\limits_{i=1,\ldots,n}\{\frac{\kappa_{ci}\varsigma_i}{2} - \frac{\lambda_i^{\max}\varsigma_i c_i}{32}\Theta_i^{*T}\Theta_i^*\}$, and $\beta_e = \sup\limits_{t\geq 0}\{\psi_e(t)\}$, (71) can be redescribed as

$$\dot{E}(t) \leq -\gamma\sum_{i=1}^{n}\|\hat{e}_i(t)\|^2 - \kappa_a\sum_{i=1}^{n}\tilde{\Theta}_{ai}^T(t)\tilde{\Theta}_{ai}(t)$$

$$-\kappa_c\sum_{i=1}^{n}\tilde{\Theta}_{ci}^T(t)\tilde{\Theta}_{ci}(t) + \beta_e. \tag{72}$$

Furthermore, according to (80) (in Remark 1), the above-mentioned inequality can be written as

$$\dot{E}(t) \leq -\frac{\gamma}{\lambda_{\max}}\hat{z}^T(t)(\tilde{L}\otimes I_m)\hat{z}(t) - \kappa_a\sum_{i=1}^{n}\tilde{\Theta}_{ai}^T(t)\tilde{\Theta}_{ai}(t)$$

$$-\kappa_c\sum_{i=1}^{n}\tilde{\Theta}_{ci}^T(t)\tilde{\Theta}_{ci}(t) + \beta_e \leq -\alpha_e E(t) + \beta_e \tag{73}$$

where $\alpha_e = \min\{\frac{2\gamma}{\lambda_{\max}}, 2\kappa_a, 2\kappa_c\}$.

According to Lemma 4, there is the fact that

$$\leq e^{-\alpha_e t}E(0) + \frac{\beta_e}{\alpha_e}\left(1 - e^{-\alpha_e t}\right)$$

From the above-mentioned inequality, it can be concluded that all error signals $z_i(t)$, $\tilde{W}_{ci}(t)$, $\tilde{W}_{ai}(t)$, $i = 1,\ldots,n$ are SGUUB.

2) Let $E_z(t) = \frac{1}{2}\hat{z}^T(t)(\tilde{L}\otimes I_m)\hat{z}(t)$, its time derivative along (40) is

$$\dot{E}_z(t) = \sum_{i=1}^{n}\left(-k_i\hat{e}_i^T(t)\tilde{x}_i(t) + \hat{e}_i^T(t)\hat{\Theta}_{fi}^T(t)\varphi_{fi}(x_i)\right.$$

$$\left. -\hat{e}_i^T(t)\dot{x}_d(t) + \hat{e}_i^T(t)u_i\right). \tag{74}$$

Performing the control (52) to the above-mentioned equation yields

$$\dot{E}_z(t) = -\sum_{i=1}^{n}\gamma_i\|\hat{e}_i(t)\|^2 + \sum_{i=1}^{n}\left(\hat{e}_i^T(t)\hat{\Theta}_{fi}^T(t)\varphi_{fi}(x_i)\right.$$

$$\left. -k_i\hat{e}_i^T(t)\tilde{x}_i(t) - \frac{1}{2}\hat{e}_i^T\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}(t) - \hat{e}_i^T(t)\dot{x}_d\right). \tag{75}$$

Applying (60) and the following inequality

$$-\frac{1}{2}\hat{e}_i^T(t)\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}(t) \leq \|\hat{e}_i(t)\|^2 + \left\|\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}(t)\right\|^2$$

to (75) has

$$\dot{E}_z(t) \leq -\gamma\|\hat{e}(t)\|^2 + \psi_z(t) \tag{76}$$

where

$$\psi_z(t) = \sum_{i=1}^{n}\left(\frac{1}{2}\|\dot{x}_d\|^2 + \frac{k_i}{4}\|\tilde{x}_i\|^2 + \frac{1}{2}\left\|\hat{\Theta}_{fi}^T(t)\varphi_{fi}(x_i)\right\|^2\right.$$

$$\left. + \left\|\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}(t)\right\|^2\right).$$

From Theorem 1 and part 1, it can be concluded that all terms of $\psi_z(t)$ are bounded. Therefore, there exists a constant β_z such that $\psi_z(t) \leq \beta_z$. Furthermore, based on (80) (in Remark 1), there is the following equation:

$$\dot{E}_z(t) \leq -\frac{\gamma}{\lambda_{\max}}\hat{z}^T(t)(\tilde{L}\otimes I_m)\hat{z}(t) + \beta_z$$

$$= -\alpha_z E_z(t) + \beta_z \tag{77}$$

where $\alpha_z = \frac{2\gamma}{\lambda_{\max}}$.

According to Lemma 4, the following result can be obtained:

$$E_z(t) \leq e^{-\alpha_z t}E_z(0) + \frac{\beta_z}{\alpha_z}\left(1 - e^{-\alpha_z t}\right). \tag{78}$$

The above-mentioned inequality implies that the tracking errors can arrive at the desired accuracy by making α_z large enough, as a result, the desired control performance can be obtained. $\square$

Remark 1: Since $\tilde{L}$ is a positive definite matrix in accordance with Lemma 2, it has n positive eigenvalues that are denoted by $\lambda_1, \lambda_2, \ldots, \lambda_n$. Let $\chi_1, \chi_2, \ldots, \chi_n$ denote the eigenvectors associated with these eigenvalues. According to matrix theory, $\chi_1, \chi_2, \ldots, \chi_n$ can be chosen to be a set of orthogonal bases. Let

$Q = [\chi_1, \chi_2, \ldots, \chi_n] \in R^{n \times n}$ and $P = \text{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$, there are the facts that $Q^T Q = QQ^T = I_n$ and $\tilde{L} = Q^T P Q$. Then the term $\hat{z}^T(t)(\tilde{L} \otimes I_m)\hat{z}(t)$ can be reexpressed as

$$\hat{z}^T(t)(\tilde{L} \otimes I_m)\hat{z}(t) = \hat{z}^T(t)\left((Q^T P Q) \otimes I_m\right)\hat{z}(t)$$

$$= \hat{z}^T(t)\left((Q^T P Q Q^T P^{-1} Q Q^T P Q) \otimes I_m\right)\hat{z}(t)$$

$$= \hat{z}^T(t)(\tilde{L} \otimes I_m)^T\left((Q^T P^{-1} Q) \otimes I_m\right)(\tilde{L} \otimes I_m)\hat{z}(t)$$

$$= \hat{e}^T(t)\left((Q^T P^{-1} Q) \otimes I_m\right)\hat{e}(t). \quad (79)$$

From the above-mentioned inequality, the following result can be yielded:

$$\lambda_{\min} \|\hat{e}(t)\|^2 \leq \hat{z}^T(t)(\tilde{L} \otimes I_m)\hat{z}(t) \leq \lambda_{\max} \|\hat{e}(t)\|^2 \quad (80)$$

where $\lambda_{\min}$ and $\lambda_{\max}$ denote the minimum and maximum eigenvalues of $Q^T P^{-1} Q$.

IV. SIMULATION EXAMPLE

In order to further demonstrate the effectiveness of the proposed formation methods, a numerical multi-agent formation consisting of four agents is carried out. In this example, the four agents move on the two-dimensional plane and the multi-agent is molded by the following dynamic:

$$\dot{x}_i(t) = -\alpha_i x_i(t) - \begin{bmatrix} 0.5 x_{i1} \cos^2(\beta_i x_{i1}) \\ x_{i2} - \sin^2(\beta_i x_{i2}) \end{bmatrix} + u_i,$$

$$i = 1, 2, 3, 4 \quad (81)$$

where $\alpha_{i=1,2,3,4} = 0.7, -3.1, 6.5, -11$ and $\beta_{i=1,2,3,4} = 0.5, 0.4, -5.5, -10$, respectively. The initial positions are $x_{i=1,2,3,4}(0) = [6, 6]^T, [-6, 6]^T, [6, -6]^T, [-6, -6]^T$, respectively.

The desired reference signal is

$$x_d(t) = [2 \sin(0.7t), 2 \cos(0.7t)]^T \quad (82)$$

of which the initial state is $x_d(0) = [-1, 1]^T$. The formation pattern is $\eta_{i=1,2,3,4} = [4; 4]^T, [-4; 4]^T, [4; -4]^T, [-4; -4]^T$.

The adjacency matrix is

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

The connection weight matrix between agents and leader is $B = \text{diag}\{1, 0, 0, 0\}$.

The identifier design: The fuzzy membership functions for agent i, $i = 1, 2, 3, 4$, are chosen as

$$\mu_{Fj}^i(x_i) = \exp\left(-\frac{\left\|x_i - [6, 6]^T + [2j-1, 2j-1]^T\right\|^2}{2}\right)$$

$$j = 1, \ldots, 6. \quad (83)$$

Then the fuzzy basis function vector is obtained as $\varphi_{fi}(x_i) = [\varphi_{fi}^1(x_i), \ldots, \varphi_{fi}^6(x_i)]$, where $\varphi_{fi}^j(x_i) = \frac{\mu_{Fj}^i(x_i)}{\sum_{j=1}^6 \mu_{Fj}^i(x_i)}$, $j = 1, \ldots, 6$. Based on (23), the adaptive identifier is built in
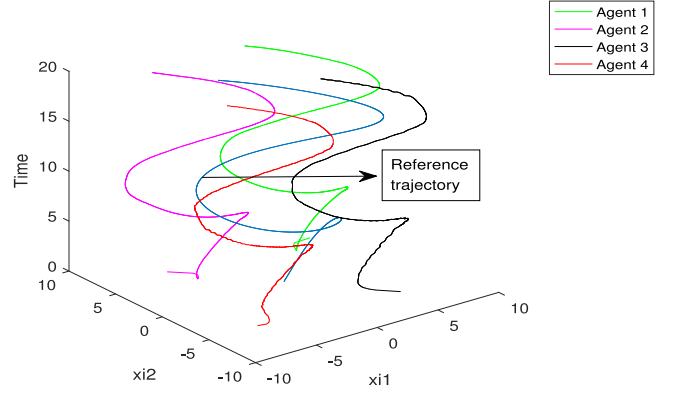

Fig. 1. Multi-agent formation performance.

the following by choosing the design parameters $k_{i=1,2,3,4} = 24, 20, 18, 16$; $\Gamma_{i=1,2,3,4} = 0.4 I_6$; and $\sigma_{i=1,2,3,4} = 0.6$:

$$\dot{\hat{x}}_i(t) = -k_i \tilde{x}_i(t) + \hat{\Theta}_{fi}^T(t)\varphi_{fi}(x_i) + u_i,$$

$$\dot{\hat{\Theta}}_{fi}(t) = 0.4\left(-\varphi_{fi}(x_i)\tilde{x}_i^T(t) - 0.6\hat{\Theta}_{fi}(t)\right) \quad (84)$$

where

$$\dot{\hat{\Theta}}_{fi}(0) = \begin{bmatrix} 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 \\ 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 \end{bmatrix}^T.$$

The optimized formation control design: The fuzzy membership functions for the distributed controller of agent i, $i = 1, 2, 3, 4$, are chosen as

$$\mu_{Fj}^i(e_i) = \exp\left(-\frac{\left\|e_i - [6, 6]^T + [2j-1, 2j-1]^T\right\|^2}{2}\right)$$

$$j = 1, \ldots, 6. \quad (85)$$

The fuzzy basis function vector is yielded as $\varphi_i(e_i) = [\varphi_i^1(e_i), \ldots, \varphi_i^6(e_i)]$, where $\varphi_i^j(e_i) = \frac{\mu_{Fj}^i(e_i)}{\sum_{j=1}^6 \mu_{Fj}^i(e_i)}$. For the actor and critic adaptive laws (55) and (56), the design parameters are chosen as $\kappa_{ai} = 0.1$ and $\kappa_{ci} = 0.2$, $i = 1, 2, 3, 4$; the initial values for adaptive adjusting vectors are $\hat{\Theta}_{ai}(0) = [0.1, 0.1, 0.1, 0.1, 0.1, 0.1]^T$ and $\hat{\Theta}_{ci}(0) = [0.2, 0.2, 0.2, 0.2, 0.2, 0.2]^T$, $i = 1, 2, 3, 4$. The control parameters are chosen as $\gamma_{i=1,2,3,4} = 26, 24, 22, 20$, respectively. According to (52), the controller is described in the following:

$$u_i = -\gamma_i \hat{e}_i(t) - \frac{1}{2}\frac{\partial^T \varphi_i(\hat{e}_i)}{\partial \hat{e}_i}\hat{\Theta}_{ai}(t), \quad i = 1, 2, 3, 4. \quad (86)$$

Simulation results are shown in Figs. 1–6. Fig. 1 displays the multi-agent formation performance. Fig. 2 shows the identifier errors. The boundedness of identifier parameter matrices, critic, and actor parameter vectors is displayed in Figs. 3–5. The cost functions are shown in Fig. 6. The simulation results further demonstrate that the proposed optimized formation scheme can guarantee the control objective to be achieved.
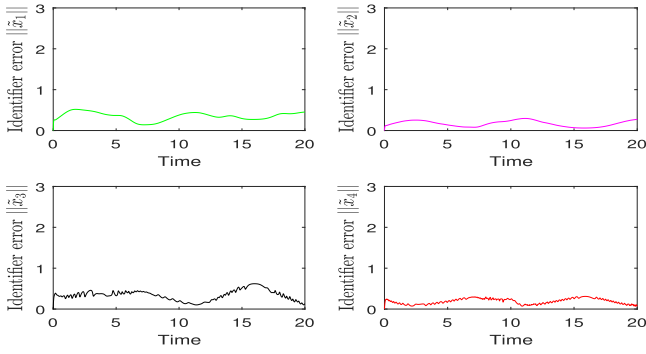
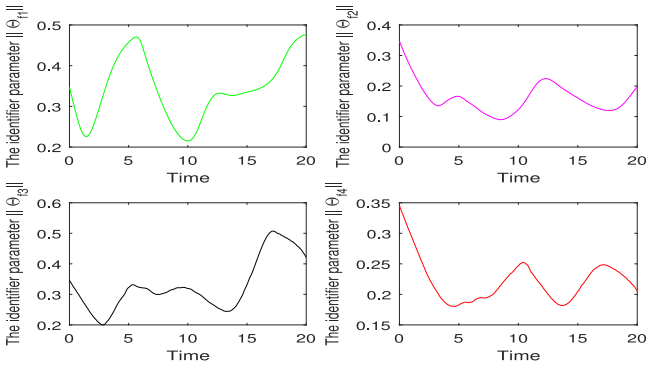Fig. 2. Norm $\|\tilde{x}_i\|$, $i = 1, 2, 3, 4$, of the adaptive identifier error.

Fig. 3. Norm $\left\|\hat{\Theta}_{fi}\right\|$, $i = 1, 2, 3, 4$, of the identifier parameter matrix.
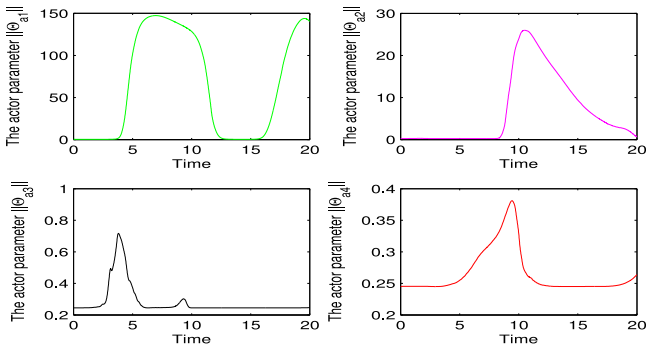

Fig. 4. Norm $\left\|\hat{\Theta}_{ai}\right\|$, $i = 1, 2, 3, 4$, of the actor parameter vector.
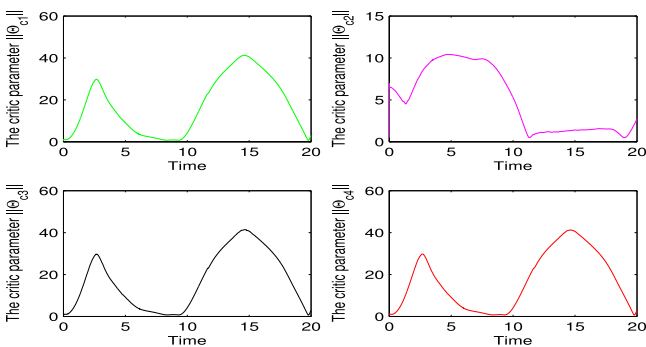

Fig. 5. Norm $\left\|\hat{\Theta}_{ci}\right\|$, $i = 1, 2, 3, 4$, of the critic parameter vector.
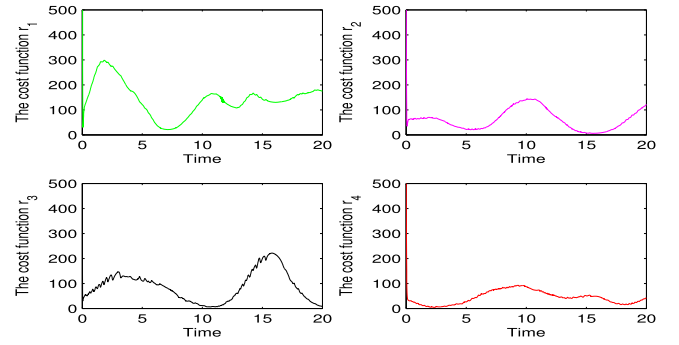

Fig. 6. Cost function $r_i(x_i, u_i)$, $i = 1, 2, 3, 4$.

V. CONCLUSION

The paper proposes an optimized control scheme for leader–follower formation of nonlinear multi-agent systems with unknown dynamics. In order to achieve the control objective, the identifier–actor–critic RL algorithm is employed based on the universal approximation property of FLS, in which the identifier is utilized to estimate the unknown dynamic of the multi-agent system; the actor FLS is utilized to carry out the control behavior; and the critic FLS is utilized to evaluate the optimizing performance and return the evaluation to the actor training. According to the Lyapunov stability theory, it is proven that the proposed scheme can achieve the control objective. Simulation results display the effectiveness of the proposed control approach.

REFERENCES

[1] L. Consolini, F. Morbidi, D. Prattichizzo, and M. Tosques, "Brief paper: Leader-follower formation control of nonholonomic mobile robots with input constraints," *Automatica*, vol. 44, no. 5, pp. 1343–1349, 2008.

[2] J. Dong, H. T. Chen, and S. Liu, "A behavior-based policy for multirobot formation control," *Appl. Mech. Mater.*, vol. 220–223, pp. 1181–1185, 2012.

[3] M. A. Lewis and K. H. Tan, "High precision formation control of mobile robots using virtual structures," *Auton. Robots*, vol. 4, no. 4, pp. 387–403, 1997.

[4] R. Olfati-Saber and R. M. Murray, "Distributed cooperative control of multiple vehicle formations using structural potential functions," *IFAC Proc. Vol.*, vol. 35, no. 1, pp. 495–500, 2002.

[5] R. E. Bellman and R. Corporation, *Dynamic Programming*. Princeton, NJ, USA: Princeton Univ. Press, 1957.

[6] L. S. Pontryagin, "Robust process control," *Uspehi Mat Nauk*, vol. 14, pp. 3–20, 1959 (in Russian).

[7] Y. Liu and Z. Geng, "Finite-time optimal formation control of multi-agent systems on the lie group SE(3)," *Int. J. Control*, vol. 86, no. 10, pp. 1675–1686, 2013.

[8] Y. Liu and Z. Geng, "Finite-time optimal formation tracking control of vehicles in horizontal plane," *Nonlinear Dyn.*, vol. 76, no. 1, pp. 481–495, 2014.

[9] W. Zhang and J. Hu, "Optimal multi-agent coordination under tree formation constraints," *IEEE Trans. Automat. Control*, vol. 53, no. 3, pp. 692–705, Apr. 2008.

[10] S. Mozer, M. C, and M. Hasselmo, "Reinforcement learning: An introduction," *IEEE Trans. Neural Netw.*, vol. 16, no. 1, pp. 285–286, Jan. 2005.

[11] A. G. Barto, R. S. Sutton, and C. W. Anderson, "Neuronlike adaptive elements that can solve difficult learning control problems," *IEEE Trans. Syst. Man Cybern.*, vol. SMC-13, no. 5, pp. 834–846, Sep./Oct. 1983.

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

[13] S. Tong, T. Wang, Y. Li, and B. Chen, "A combined backstepping and stochastic small-gain approach to robust adaptive fuzzy output feedback control," *IEEE Trans. Fuzzy Syst.*, vol. 21, no. 2, pp. 314–327, Apr. 2013.

[14] R. E. Precup, R. C. David, A. I. Szedlak-Stinean, E. M. Petriu, and F. Dragan, "An easily understandable grey wolf optimizer and its application to fuzzy controller tuning," *Algorithms*, vol. 10, no. 2, 2017, Art. no. 68.

[15] Q. Zhou, H. Li, C. Wu, L. Wang, and C. K. Ahn, "Adaptive fuzzy control of nonlinear systems with unmodeled dynamics and input saturation using small-gain approach," *IEEE Trans. Syst. Man Cybern. Syst.*, vol. 47, no. 8, pp. 1979–1989, Aug. 2017.

[16] H. Li, L. Bai, L. Wang, Q. Zhou, and H. Wang, "Adaptive neural control of uncertain nonstrict-feedback stochastic nonlinear systems with output constraint and unknown dead zone," *IEEE Trans. Syst. Man Cybern. Syst.*, vol. 47, no. 8, pp. 2048–2059, Aug. 2017.

[17] R. E. Precup, C. A. Bojan-Dragos, E. L. Hedrea, I. D. Borlea, and E. M. Petriu, "Evolving fuzzy models for anti-lock braking systems," in *Proc. IEEE Int. Conf. Comput. Intell. Virtual Environ. Meas. Syst. Appl.*, 2017, pp. 48–53.

[18] O. Cordon, A. Peregrin, and F. Herrera, "Searching for basic properties obtaining robust implication operators in fuzzy control," *Fuzzy Sets Syst.*, vol. 111, no. 2, pp. 237–251, 2000.

[19] R. E. Precup, R. C. David, and E. M. Petriu, "Grey wolf optimizer algorithm-based tuning of fuzzy control systems with reduced parametric sensitivity," *IEEE Trans. Ind. Electron.*, vol. 64, no. 1, pp. 527–534, Jan. 2017.

[20] Q. Zhou, H. Li, L. Wang, and R. Lu, "Prescribed performance observer-based adaptive fuzzy control for nonstrict-feedback stochastic nonlinear systems," *IEEE Trans. Syst. Man Cybern., Syst.*, 2017, doi: 10.1109/TSMC.2017.2738155.

[21] R. E. Precup, M. L. Tomescu, S. Preitl, and E. M. Petriu, "Fuzzy logic-based stabilization of nonlinear time-varying systems," *Int. J. Artif. Intell.*, vol. 3, no. 9, pp. 24–36, 2008.

[22] L. Wang, M. Basin, H. Li, and R. Lu, "Observer-based composite adaptive fuzzy control for nonstrict-feedback systems with actuator failures," *IEEE Trans. Fuzzy Syst.*, 2017, doi: 10.1109/TFUZZ.2017.2774185.

[23] R. E. Precup, M. L. Tomescu, S. Preitl, E. M. Petriu, J. Fodor, and C. Pozna, "Stability analysis and design of a class of MIMO fuzzy control systems," *J. Intell. Fuzzy Syst.*, vol. 25, no. 1, pp. 145–155, 2013.

[24] H. L. Dong, Y. H. Joo, and M. H. Tak, "Local analysis of continuous-time Takagi-Sugeno fuzzy systems: A fuzzy Lyapunov function approach," *Inf. Sci.*, vol. 257, no. 2, pp. 163–175, 2014.

[25] R. E. Precup, E. M. Petriu, D. Claudia-Adina, and D. Radu-codru, "Stability analysis results concerning the fuzzy control of a class of nonlinear time-varying systems," *Theory Appl. Math. Comput. Sci.*, vol. 1, no. 1, pp. 2–10, 2011.

[26] J. Peng, M. Liu, J. Liu, K. C. Lin, and M. Wu, "A coordination model using fuzzy reinforcement learning for multi-agent system," in *Proc. Int. Symp. Collaborative Technol. Syst.*, 2008, pp. 129–136.

[27] D. Luviano and W. Yu, "Continuous-time path planning for multi-agents with fuzzy reinforcement learning," *J. Intell. Fuzzy Syst.*, vol. 33, no. 1, pp. 491–501, 2017.

[28] K. G. Vamvoudakis and F. L. Lewis, "Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem," *Automatica*, vol. 46, no. 5, pp. 878–888, 2010.

[29] L. X. Wang, "Fuzzy systems are universal approximators," in *Proc. IEEE Int. Conf. Fuzzy Syst.*, 1992, pp. 1163–1170.

[30] G. Wen, C. L. P. Chen, Y. J. Liu, and L. Zhi, "Neural network-based adaptive leader-following consensus control for a class of nonlinear multiagent state-delay systems," *IEEE Trans. Cybern.*, vol. 47, no. 8, pp. 2151–2160, Aug. 2017.

[31] J.-L. Wang and H.-N. Wu, "Leader-following formation control of multi-agent systems under fixed and switching topologies," *Int. J. Control*, vol. 85, no. 6, pp. 695–705, 2012.

[32] R. W. Beard, G. N. Saridis, and J. T. Wen, "Galerkin approximations of the generalized Hamilton-Jacobi-Bellman equation," *Automatica*, vol. 33, no. 12, pp. 2159–2177, 1997.

[33] J. M. Aldaz, S. Barza, M. Fujii, and M. S. Moslehian, "Advances in operator Cauchy–Schwarz inequalities and their reverses," *Ann. Funct. Anal.*, vol. 6, no. 3, pp. 275–295, 2015.

[34] G. H. Hardy, J. E. Littlewood, and G. Plya, *Inequalities (Cambridge Mathematical Library)*. Cambridge, U.K.: Cambridge Univ. Press, 1934.

Guoxing Wen received the M.S. degree in applied mathematics from Liaoning University of Technology, Jinzhou, China, in 2011, and the Ph.D. degree in computer and information science from Macau University, Macau, China, in 2014.

He was a Research Fellow with the Department of Electrical and Computer Engineering, Faculty of Engineering, National University of Singapore, Singapore, from Sept. 2015 to Sept. 2016. He is currently a Lecturer with the College of Science, Binzhou University, Binzhou, China. His research interests include adaptive neural network control, optimal control, and multi-agent control.

C. L. Philip Chen (S'88–M'88–SM'94–F'07) received the M.S. degree from the University of Michigan, Ann Arbor, MI, USA, and the Ph.D. degree from Purdue University, West Lafayette, IN, USA, in 1985 and 1988, respectively, both in electrical engineering.

He is currently a Chair Professor with the Department of Computer and Information Science and the Dean of the Faculty of Science and Technology, University of Macau, Macau, China. His current research interests include computational intelligence, systems, and cybernetics.

Jun Feng received the M.S. degree in computer applied technology from Qufu Normal University, Qufu, China, in 2009. She is currently working toward the Ph.D. degree at Nanjing Aeronautics and Astronautics University, Nanjing, China.

She is also an Associate Professor with the Department of Information Engineering, Binzhou University, Binzhou, China. Her research interests include consensus control of nonlinear multi-agent systems and artificial intelligence.

Ning Zhou received the M.S. degree in applied mathematics from Liaoning University of Technology, Jinzhou, China, in 2011, and the Ph.D. degree in control science and engineering from Beijing Institute of Technology, Beijing, China, in 2015.

She is currently a Postdoctoral Scholar with the Faculty of Science and Engineering, University of Groningen, Groningen, The Netherlands, and also a Lecturer with the Research Institute of Big Data for Agriculture and Forestry, College of Computer and Information Sciences, Fujian Agriculture and Forestry University, Fuzhou, China. Her research interests include spacecraft attitude synchronization, sliding mode control, multi-agent control, and fuzzy control.