专家情况登记表

专家基本信息：

姓名	文国兴	性别	男	籍贯	山东滨州	
民族	汉族	政治面貌		群众		
出生日期	1977-02-15	身份证号		37230119770215075X		
党内职务	无	申报类型		一般推荐		
最高学历	博士研究生			最高学位	博士	
毕业院校	澳门大学			所学专业	软件工程	
从事专业所属一级学科	工学			从事专业所属二级学科	控制科学与工程	
是否为留学归国人员				是否为博士后	是	
参加工作时间	2016-01-01			职业	教学人员	
聘任岗位	教师			聘任时间	2019-05-01	
行政职务	无			行政级别	无	
专业技术职务	高等学校教师-教授			联系方式	办公电话：0543 3191176 手机：15266765110	
通讯地址	山东省滨州市滨城区黄河五路391号滨州学院理学院			邮编	256600	
邮箱	wengx_bzu@hotmail.com					
入选“百千万人才工程”国家级人选时间						
入选“享受国务院颁发政府特殊津贴”人员时间						

教育经历：

起始时间	结束时间	学校	专业	学历	学位
2004-01	2006-01	山东教育学院	数学与应用数学	大学本科	学士
2008-09	2011-04	辽宁工业大学	应用数学	硕士研究生	硕士
2011-09	2014-11	澳门大学	软件工程	博士研究生	博士

工作经历：

起始年度	终止年度	工作单位	部门	行政职务	专业技术职务
1997-09	2011-09	滨州市梁才乡教委	梁才乡教委	无	小教二级
2015-09	2016-09	新加坡国立大学	工程学院	无	博士后
2016-09	2019-07	滨州学院	理学院	无	校聘教授
2018-06	2018-09	澳门大学	科技学院	无	博士后

学术团体任职情况：

起始时间	结束时间	工作单位名称	职务	备注
2016-07-01	2021-07-01	中国自动化学会	委员	

重要社会兼职情况：

起始时间	结束时间	工作单位名称	职务	备注
2018-01-01	2021-12-31	曲阜师范大学	研究生导师	

代表性著作和论文情况：

著作或论文题目	出版或发表时间	收录情况或出版社名称	影响因子	是否为通讯作者	位次/人数
Artificial Potential Based Adaptive H∞ Synchronized Tracking Control for Accommodation Vessel	2017-07-01	IEEE Transactions on Industrial Electronics	7.503 (2018年度)	是	1/3
Neural-Network-Based Adaptive Leader-Following Consensus Control for a Class of Nonlinear Multi-Agent State-Delay Systems	2017-08-01	IEEE Transactions on Cybernetics	10.387 (2018年度)	是	1/4

Formation Control with Obstacle Avoidance for a class of Stochastic Multi-Agent Systems	2018-07-01	IEEE Transactions on Industrial Electronics	7.503 (2018年度)	是	1/3
Optimized Backstepping for Tracking Control of Strict Feedback Systems	2018-08-01	IEEE Transactions on neural networks and learning systems	11.683 (2018年度)	是	1/3
Optimized Multi-Agent Formation Control Based on Identifier-Actor-Critic Reinforcement Learning Algorithm	2018-10-01	IEEE Transactions on Fuzzy Systems	8.759 (2018年度)	是	1/4

成果获奖情况：

年度	获奖种类	获奖项目情况	等级	位次/人数	发证单位名称	备注
2017	科学技术进步奖	SCI文章获奖	一等奖	1/3	滨州市委组织部	
2018	其他	SCI文章获奖	其他	2/4	IEEE Computational Intelligence Society	

个人荣誉情况及获选人才工程情况：

起始时间	终止时间	荣誉名称或工程名称	授予单位或主管部门	工程支持资金总额（人民币万元）	层级	备注
2018-01-01	2021-12-31	聚英计划	滨州学院	7.2万	市级	

主持的项目课题情况：

结题时间	项目、课题名称	项目、课题类别	位次/总人数	备注
2021-06-01	多智能体编队的优化控制	省（部）级	1/7	在研

已授权专利情况：

专利名称	专利类别	专利号	获批时间	备注

业绩贡献：

主要业绩贡献
文国兴，男，汉族，1977年2月出生，滨州学院校聘教授；澳门大学博士研究生，师从陈俊龙教授；新加坡国立大学博士后，师从新加坡工程院士葛树志教授。自2015年入职滨州学院以来，主要科研成果主要包括：（1）主持山东省自然科学基金面上项目1项。（2）发表SCI学术论文共12篇，其中有3篇全球ESI高被引论文（2篇为一作和通讯，1篇为二作），有9篇一作文章，并7篇兼通讯作者发表在IEEE Trans系列期刊（理工类权威期刊，中科院一区）。（3）获得滨州市自然科学优秀学术成果奖一等奖1次。（4）入选滨州学院“聚英计划”人才计划。（5）2017年，本人博士期间论文 Adaptive Consensus Control for a Class of Nonlinear Multi-agent Time-Delay Systems Using Neural Networks（导师陈俊教授一作，本人二作），被本专业权威期刊IEEE Transactions on neural networks and learning systems授于Outstanding Paper Award。

单位联系人信息:

单位名称	**滨州学院**	**联系人所在部门**	**人事处**
单位人事联系人	**张志彬**	**邮箱**	bzurcb@163.com
办公电话	0543-3191176	**手机**	15954310661
传真	0543-3190000	**单位所在区域**	**山东省滨州市滨城区**
邮编	256603	**通讯地址**	**山东省学州市黄河五路391号滨州学院**

附件列表

序号	附件类型	附件名称
1	代表论文、论著	paper 1.pdf
2	代表论文、论著	paper1proof.pdf
3	代表论文、论著	paper 2.pdf
4	代表论文、论著	paper2proof.pdf
5	代表论文、论著	paper 3.pdf
6	代表论文、论著	paper3proof.pdf
7	代表论文、论著	paper 4.pdf
8	代表论文、论著	paper4proof.pdf
9	代表论文、论著	paper 5.pdf
10	代表论文、论著	paper5proof.pdf
11	个人承诺协议书	承诺书.jpg
12	工作单位证明材料	教师证明.jpg
13	工作经历信息	工作经历证明-文国兴.pdf
14	工作经历信息	工作经历证明-文国兴.pdf
15	工作经历信息	6 .新加坡博士后留学证明.pdf
16	工作经历信息	新加坡国立博士后证书.pdf
17	工作经历信息	澳门政府批示.pdf
18	工作经历信息	2018.06 澳门大学合同.pdf
19	个人荣誉情况及获选人才工程情况	聚英计划.pdf
20	用户获奖信息	2017滨州市自然科学一等奖.pdf
21	用户获奖信息	TNNLS杰出论文Chen.pdf
22	教育经历	5 .澳门大学博士毕业证.pdf
23	教育经历	5 .博士学位认证.pdf
24	“ 六公开 ” 监督卡	六公开卡.jpg

25	聘任岗位	教师证明.jpg
26	用户身份证信息	身份证.pdf
27	重要社会兼职信息	曲师大关于公布2018年研究生导师考核和增选结果的通知.pdf
28	用户头像信息	白底照片.jpg
29	项目、课题信息	省基金任务书_盖章.pdf
30	学术团体任职信息	中国自动化学会委员.jpg
31	专业技术职务	教授聘任文件.jpg.jpg
32	专业技术职务	教授聘书.jpg.jpg

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).**

Manuscript received July 30, 2017; revised October 19, 2017 and December 18, 2017; accepted December 19, 2017. Date of publication December 27, 2017; date of current version October 4, 2018. This work was supported in part by the Doctoral Scientific Research Staring Fund of Binzhou University under Grant 2016Y14, in part by the National Natural Science Foundation of China under Grants 61572540, 61603094 and 61603095, in part by the China Scholarship Council under Grant 201707870005, in part by the Macau Science and Technology Development Fund under Grant 019/2015/A, Grant 024/2015/AMJ, and Grant 079/2017/A2, and in part by the University Macau MYR Grants. *(Corresponding author: Guoxing Wen.)*

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 the 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).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TFUZZ.2017.2787561

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 difficultly 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 $\theta_j = \max_{y\in R} \mu_{G^j}(y)$.

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

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

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

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

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 $\Upsilon = \{\upsilon_1, \upsilon_2, \ldots, \upsilon_n\}$, $\Xi \subseteq \Upsilon \times \Upsilon$ and $A = [a_{ij}]$ are the node set, edge set, and adjacency matrix, respectively. Let $\xi_{ij} = (\upsilon_i, \upsilon_j)$ 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 $\Lambda_i = \{\upsilon_j \| (\upsilon_i, \upsilon_j) \in \Xi\}$. 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

$$L = D - A \quad (6)$$

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 $q_{ij} = q_{ji} \le 0$ for $i \ne j$ 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 $\dot{\Phi}(t) \le -\alpha\Phi(t) + \beta$ is held, where α and β are positive constants, then there is the following result:

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

III. Main Results

A. Problem Formulation

Consider the multi-agent system modeled in the following:

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

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

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

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:

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

Define the formation errors as

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

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:

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

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

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

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

Define the infinite horizon value function as

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

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 $e^T(t) = [e_1^T, \ldots, e_n^T]$; $u = \left[u_1^T, \ldots, u_n^T\right]^T$; $z = \left[z_1^T, \ldots, z_n^T\right]^T$; $C = \mathrm{diag}\{c_1, \ldots, c_n\}$; 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 $V_i(e_i) = \int_t^\infty r_i\big(e_i(\tau), u_i(e_i)\big)d\tau$, the value function (14) can be re-expressed as

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

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

$$\begin{aligned} H\left(e, u, \frac{\partial V}{\partial e}\right) &= r(e,u) + \frac{\partial V(e)}{\partial e^T}\dot{e}(t) \\ &= e^T e + u^T(C \otimes I_m)u + \sum_{i=1}^n \left(\frac{\partial V_i(e_i)}{\partial e_i^T}\dot{e}_i(t)\right) \\ &= \sum_{i=1}^n \left(\|e_i(t)\|^2 + c_i\|u_i\|^2 + \frac{\partial V_i(e_i)}{\partial e_i^T}\dot{e}_i(t)\right) \end{aligned} \quad (16)$$

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

$$\begin{aligned} V^*(e) &= \min_{u_{i=1,\cdots,n} \in \Psi(\Omega)} \int_t^\infty r(e,u)\,d\tau = \int_t^\infty r(e,u^*)\,d\tau \\ &= \sum_{i=1}^n V_i^*(e_i) = \sum_{i=1}^n \min_{u_i \in \Psi(\Omega)} \int_t^\infty r_i(e_i,u_i)\,d\tau \\ &= \sum_{i=1}^n \int_t^\infty r_i(e_i, u_i^*)\,d\tau \end{aligned} \quad (17)$$

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

$$\begin{aligned} H\left(e, u^*, \frac{\partial V^*}{\partial e}\right) &= r(e, u^*) + \frac{\partial V^*(e)}{\partial e^T}\dot{e}(t) \\ &= \sum_{i=1}^n \left(\|e_i\|^2 + c_i\|u_i^*\|^2 + \frac{\partial V_i^*(e_i)}{\partial e_i^T}\dot{e}_i(t)\right) = 0. \end{aligned} \quad (18)$$

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

$$\begin{aligned} H_i\left(e_i, u_i^*, \frac{\partial V_i^*}{\partial e_i}\right) &= \|e_i\|^2 + c_i\|u_i^*\|^2 + \frac{\partial V_i^*(e_i)}{\partial e_i^T}\Bigg(c_i f_i(x_i) \\ &\quad + c_i u_i^* - b_i\dot{x}_d(t) - \sum_{j\in\Lambda_i} a_{ij}\dot{x}_j(t)\Bigg) = 0, \quad i = 1, \ldots, n. \end{aligned} \quad (19)$$

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

$$\begin{aligned} &\|e_i(t)\|^2 + \frac{\partial V_i^*}{\partial e_i^T}\left(c_i f_i(x_i) - b_i\dot{x}_d(t) - \sum_{j\in\Lambda_i} a_{ij}\dot{x}_j(t)\right) \\ &- \frac{c_i}{4}\frac{\partial V_i^*}{\partial e_i^T}\frac{\partial V_i^*}{\partial e_i} = 0, \quad i = 1, \ldots, n. \end{aligned} \quad (21)$$

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 $f_i(x_i), \quad i = 1, \ldots, n$, 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 $\|\varepsilon_{fi}(x_i)\| \le \delta_{fi}$, 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

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

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

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

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

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

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 $k_i,\quad i = 1,\ldots,n$, large enough.

Proof: 1) Consider the Lyapunov candidate as following:

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

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

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

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

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

Substituting (28) into (27), we obtain

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

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

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

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

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

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

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

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

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

$$\dot{E}_1(t) \le -\alpha_1E_1(t) + \beta_1. \tag{33}$$

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

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

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

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

Inserting the following facts:

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

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

to (35) yields

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

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 $\tilde{\Theta}_{f1}^T(t),\ldots,\tilde{\Theta}_{fn}^T(t)$ are bounded, which are proven by part 1, the term $\psi_x(t)$ is bounded. Let $\alpha_2 = \min_{i=1,\ldots,n}\{k_i - 1\}$ and $\beta_2 = \sup_{t\ge0}\{\psi_x(t)\}$, (36) becomes

$$\dot{E}_x(t) \le -\alpha_2E_x(t) + \beta_2. \tag{37}$$

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

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

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

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:

$$\dot{\hat{z}}_i(t) = -k_i\tilde{x}_i(t) + \hat{\Theta}_{fi}^T(t)\varphi_{fi}(x_i) - \dot{x}_d(t) + u_i, \quad (40)$$
$$\dot{\hat{e}}_i(t) = -k_ic_i\tilde{x}_i(t) + c_i\hat{\Theta}_{fi}^T(t)\varphi_{fi}(x_i) + c_iu_i - b_i\dot{x}_d - \sum_{j\in\Lambda_i} a_{ij}\dot{\hat{x}}_j(t), \quad i = 1,\ldots,n. \quad (41)$$

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

$$V^*(\hat{e}) = \min_{u_{i=1,\cdots,n}\in\Psi(\Omega)} \int_t^\infty r\left(\hat{e}(\tau), u(\hat{e})\right)d\tau = \sum_{i=1}^n V_i^*(\hat{e}_i) = \sum_{i=1}^n \min_{u_i\in\Psi(\Omega)} \int_t^\infty r_i\left(\hat{e}_i(\tau), u_i(\hat{e}_i)\right)d\tau = \sum_{i=1}^n \int_t^\infty r_i\left(\hat{e}_i(\tau), u_i^*(\hat{e}_i)\right)d\tau \quad (42)$$

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

$$H_i\left(\hat{e}_i, u_i^*, \frac{\partial V_i^*}{\partial \hat{e}_i}\right) = \|\hat{e}_i(t)\|^2 + c_i\|u_i^*\|^2 + \frac{\partial V_i^*(\hat{e}_i)}{\partial \hat{e}_i^T}\dot{\hat{e}}_i = \|\hat{e}_i(t)\|^2 + c_i\|u_i^*\|^2 + \frac{\partial V_i^*(\hat{e}_i)}{\partial \hat{e}_i^T}\Big(-k_ic_i\tilde{x}_i(t) + c_iu_i^* + c_i\hat{\Theta}_{fi}^T(t)\varphi_{fi}(x_i) - b_i\dot{x}_d - \sum_{j\in\Lambda_i} a_{ij}\dot{\hat{x}}_j(t)\Big) = 0, \quad i = 1,\ldots,n. \quad (43)$$

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

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

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

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

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

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

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

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

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 $|\varepsilon_i(\hat{e}_i)| \le \delta_i$ 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

$$V_i^*(\hat{e}_i) = \gamma_i\|\hat{e}_i(t)\|^2 + \Theta_i^{*T}\varphi_i(\hat{e}_i) + \varepsilon_i(\hat{e}_i), \quad (48)$$
$$u_i^* = -\gamma_i\hat{e}_i(t) - \frac{1}{2}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial \hat{e}_i}\Theta_i^* - \frac{1}{2}\frac{\partial\varepsilon_i(\hat{e}_i)}{\partial \hat{e}_i}, \quad i = 1,\ldots,n \quad (49)$$

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:

$$H_i\left(\hat{e}_i, u_i^*, \frac{\partial V_i^*}{\partial \hat{e}_i}\right) = -(\gamma_i^2c_i - 1)\|\hat{e}_i(t)\|^2 + 2\gamma_i\hat{e}_i^T(t) \times \left(c_i\hat{\Theta}_{fi}^T\varphi_{fi}(x_i) - k_ic_i\tilde{x}_i(t) - b_i\dot{x}_d - \sum_{j\in\Lambda_i} a_{ij}\dot{\hat{x}}_j(t)\right) + \Theta_i^{*T}\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\Big(c_i\hat{\Theta}_{fi}^T\varphi_{fi}(x_i) - \gamma_ic_i\hat{e}_i(t) - k_ic_i\tilde{x}_i(t) - b_i\dot{x}_d(t) - \sum_{j\in\Lambda_i} a_{ij}\dot{\hat{x}}_j(t)\Big) - \frac{c_i}{4}\left\|\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\Theta_i^*\right\|^2 + \epsilon_i(t) = 0 \quad (50)$$

where

$$\epsilon_i(t) = \frac{\partial\varepsilon_i(\hat{e}_i)}{\partial\hat{e}_i^T}\Big(c_iu_i^* - k_ic_i\tilde{x}_i(t) + c_i\hat{\Theta}_{fi}^T\varphi_{fi}(x_i) - b_i\dot{x}_d - \sum_{j\in\Lambda_i} a_{ij}\dot{\hat{x}}_j(t)\Big) + \frac{c_i}{4}\left\|\frac{\partial\varepsilon_i(\hat{e}_i)}{\partial\hat{e}_i}\right\|^2.$$

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:

$$\hat{V}_i^*(\hat{e}_i) = \gamma_i\|\hat{e}_i(t)\|^2 + \hat{\Theta}_{ci}^T(t)\varphi_i(\hat{e}_i), \quad (51)$$
$$u_i = -\gamma_i\hat{e}_i(t) - \frac{1}{2}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}(t), \quad i = 1,\ldots,n \quad (52)$$

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

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

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:

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

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

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

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 $\xi_i(t)\xi_i^T(t), \quad i = 1, 2, \ldots, n$, 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:

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

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

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

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

$$\begin{aligned}
&-k_i\hat{e}_i^T(t)\tilde{x}_i(t) \le k_i\|\hat{e}_i(t)\|^2 + \frac{k_i}{4}\|\tilde{x}_i(t)\|^2,\\
&\hat{e}_i^T(t)\hat{\Theta}_{fi}^T(t)\varphi_{fi}(x_i) \le \frac{1}{2}\|\hat{e}_i(t)\|^2 + \frac{1}{2}\left\|\hat{\Theta}_{fi}^T(t)\varphi_{fi}\right\|^2,\\
&-\hat{e}_i^T(t)\dot{x}_d(t) \le \frac{1}{2}\|\hat{e}_i(t)\|^2 + \frac{1}{2}\|\dot{x}_d(t)\|^2. \qquad (60)
\end{aligned}$$

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

$$\begin{aligned}
\dot{E}(t) \le& -\sum_{i=1}^n(\gamma_i - k_i - 1)\|\hat{e}_i\|^2 + \sum_{i=1}^n\Bigg(-\frac{1}{2}\hat{e}_i^T\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}\\
&+\frac{1}{2}\tilde{\Theta}_{ai}^T\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\hat{e}_i - \kappa_{ai}c_i\tilde{\Theta}_{ai}^T(t)\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}(t)\\
&+\frac{\kappa_{ci}c_i}{4\left(1+\|\xi_i\|^2\right)}\tilde{\Theta}_{ai}^T(t)\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}(t)\xi_i^T\hat{\Theta}_{ci}(t)\Bigg)\\
&+\sum_{i=1}^n\tilde{\Theta}_{ci}^T(t)\Bigg(-\frac{\kappa_{ci}\xi_i}{1+\|\xi_i\|^2}\Bigg(\xi_i^T\hat{\Theta}_{ci}(t) - (\gamma_i^2c_i - 1)\|\hat{e}_i\|^2\\
&+2\gamma_i\hat{e}_i^T\Bigg(\hat{\Theta}_{fi}^T(t)\varphi_{fi}(x_i) - k_ic_i\tilde{x}_i - b_i\dot{x}_d - \sum_{j\in\Lambda_i}a_{ij}\dot{\hat{x}}_j\Bigg)\\
&+\frac{c_i}{4}\left\|\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}(t)\right\|^2\Bigg)\Bigg) + \sum_{i=1}^n\Bigg(\frac{k_i}{4}\|\tilde{x}_i\|^2 + \frac{1}{2}\|\dot{x}_d\|^2\\
&+\frac{1}{2}\left\|\hat{\Theta}_{fi}^T(t)\varphi_{fi}(x_i)\right\|^2\Bigg). \qquad (61)
\end{aligned}$$

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

$$\begin{aligned}
&\tilde{\Theta}_{ai}^T(t)\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\hat{e}_i - \hat{e}_i^T(t)\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ci} = -\hat{e}_i^T\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\Theta_i^*,\\
&-\kappa_{ai}c_i\tilde{\Theta}_{ai}^T(t)\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}(t) = -\frac{\kappa_{ai}c_i}{2}\tilde{\Theta}_{ai}^T(t)\\
&\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\tilde{\Theta}_{ai}(t) - \frac{\kappa_{ai}c_i}{2}\hat{\Theta}_{ai}^T(t)\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\\
&\hat{\Theta}_{ai}(t) + \frac{\kappa_{ai}c_i}{2}\Theta_i^{*T}\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\Theta_i^*.
\end{aligned}$$

Substituting the above-mentioned equations into (61) yields

$$\begin{aligned}
\dot{E}(t) \le& -\sum_{i=1}^n(\gamma_i - k_i - 1)\|\hat{e}_i\|^2 - \frac{1}{2}\sum_{i=1}^n\hat{e}_i^T\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\Theta_i^*\\
&-\sum_{i=1}^n\frac{\kappa_{ai}c_i}{2}\tilde{\Theta}_{ai}^T\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\tilde{\Theta}_{ai} - \sum_{i=1}^n\frac{\kappa_{ai}c_i}{2}\hat{\Theta}_{ai}^T(t)\\
&\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}(t) + \sum_{i=1}^n\frac{\kappa_{ci}c_i}{4\left(1+\|\xi_i\|^2\right)}\\
&\tilde{\Theta}_{ai}^T(t)\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}(t)\xi_i^T(t)\hat{\Theta}_{ci}(t) + \sum_{i=1}^n\tilde{\Theta}_{ci}^T(t)\\
&\Bigg(-\frac{\kappa_{ci}\xi_i(t)}{1+\|\xi_i(t)\|^2}\Big(\xi_i^T\hat{\Theta}_{ci}(t) - (\gamma_i^2c_i - 1)\|\hat{e}_i\|^2 + 2\gamma_i\hat{e}_i^T(t)\\
&\Bigg(\hat{\Theta}_{fi}^T(t)\varphi_{fi}(x_i) - k_ic_i\tilde{x}_i - b_i\dot{x}_d - \sum_{j\in\Lambda_i}a_{ij}\dot{\hat{x}}_j\Bigg)\\
&+\frac{c_i}{4}\left\|\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}(t)\right\|^2\Bigg) + \sum_{i=1}^n\Bigg(\frac{k_i}{4}\|\tilde{x}_i(t)\|^2 + \frac{1}{2}\|\dot{x}_d(t)\|^2\\
&+\frac{1}{2}\left\|\hat{\Theta}_{fi}^T(t)\varphi_{fi}(x_i)\right\|^2 + \frac{\kappa_{ai}c_i}{2}\Theta_i^{*T}\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\Theta_i^*\Bigg). \qquad (62)
\end{aligned}$$

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

$$\begin{aligned}
&-(\gamma_i^2c_i - 1)\|\hat{e}_i(t)\|^2 + 2\gamma_i\hat{e}_i^T(t)\Big(c_i\hat{\Theta}_{fi}^T(t)\varphi_{fi}(x_i) - k_ic_i\tilde{x}_i\\
&-b_i\dot{x}_d(t) - \sum_{j\in\Lambda_i}a_{ij}\dot{\hat{x}}_j(t)\Bigg) = -\xi_i^T(t)\Theta_i^* - \frac{c_i}{2}\Theta_i^{*T}\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\\
&\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}(t) + \frac{c_i}{4}\left\|\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\Theta_i^*\right\|^2 - \epsilon_i(t). \qquad (63)
\end{aligned}$$

Applying (63) and the fact that

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

(62) can be rewritten as

$$\begin{aligned}
\dot{E}(t) \le& -\sum_{i=1}^n(\gamma_i - k_i - 2)\|\hat{e}_i\|^2 - \sum_{i=1}^n\frac{\kappa_{ai}c_i}{2}\tilde{\Theta}_{ai}^T\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\\
&\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\tilde{\Theta}_{ai} - \sum_{i=1}^n\frac{\kappa_{ai}c_i}{2}\hat{\Theta}_{ai}^T\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}\\
&+\sum_{i=1}^n\frac{\kappa_{ci}c_i}{4\left(1+\|\xi_i\|^2\right)}\tilde{\Theta}_{ai}^T(t)\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}(t)\\
&\xi_i^T(t)\hat{\Theta}_{ci}(t) + \sum_{i=1}^n\tilde{\Theta}_{ci}^T(t)\Bigg(-\frac{\kappa_{ci}\xi_i(t)}{1+\|\xi_i(t)\|^2}\Big(\xi_i^T(t)\tilde{\Theta}_{ci}(t)\\
&+\frac{c_i}{4}\left\|\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}(t)\right\|^2 - \frac{c_i}{2}\Theta_i^{*T}\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\\
&\hat{\Theta}_{ai}(t) + \frac{c_i}{4}\left\|\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\Theta_i^*\right\|^2 - \epsilon_i\Big)\Bigg) + \sum_{i=1}^n\Bigg(\frac{k_i}{4}\|\tilde{x}_i\|^2
\end{aligned}$$

$$+\frac{1}{2}\left\|\hat{\Theta}_{fi}^T(t)\varphi_{fi}(x_i)\right\|^2+\frac{\kappa_{ai}c_i+2}{2}\left\|\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\Theta_i^*\right\|^2$$
$$+\frac{1}{2}\|\dot{x}_d\|^2\Bigg). \quad (65)$$

Using the fact that

$$\frac{c_i}{4}\left\|\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}(t)\right\|^2-\frac{c_i}{2}\Theta_i^{*T}\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}(t)$$
$$+\frac{c_i}{4}\left\|\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\Theta_i^*\right\|^2=\frac{c_i}{4}\tilde{\Theta}_{ai}^T(t)\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}(t)$$
$$-\frac{c_i}{4}\Theta_i^{*T}\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\tilde{\Theta}_{ai}(t) \quad (66)$$

(65) can be rewritten as

$$\dot{E}(t)\le-\sum_{i=1}^{n}(\gamma_i-k_i-2)\|\hat{e}_i(t)\|^2-\sum_{i=1}^{n}\frac{\kappa_{ai}c_i}{2}\tilde{\Theta}_{ai}^T(t)$$
$$\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\tilde{\Theta}_{ai}(t)-\sum_{i=1}^{n}\frac{\kappa_{ci}}{1+\|\xi_i\|^2}\tilde{\Theta}_{ci}^T(t)\xi_i\xi_i^T\tilde{\Theta}_{ci}(t)$$
$$+\sum_{i=1}^{n}\frac{\kappa_{ci}c_i}{4\left(1+\|\xi_i\|^2\right)}\tilde{\Theta}_{ai}^T(t)\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}(t)\xi_i^T\tilde{\Theta}_{ci}$$
$$-\sum_{i=1}^{n}\frac{c_i\kappa_{ci}}{4\left(1+\|\xi_i\|^2\right)}\tilde{\Theta}_{ci}^T(t)\xi_i\tilde{\Theta}_{ai}^T(t)\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}$$
$$+\sum_{i=1}^{n}\frac{c_i\kappa_{ci}}{4\left(1+\|\xi_i\|^2\right)}\tilde{\Theta}_{ci}^T(t)\xi_i\Theta_i^{*T}\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\tilde{\Theta}_{ai}(t)$$
$$+\sum_{i=1}^{n}\frac{\kappa_{ci}}{1+\|\xi_i\|^2}\tilde{\Theta}_{ci}^T(t)\xi_i\epsilon_i(t)-\sum_{i=1}^{n}\frac{\kappa_{ai}c_i}{2}\hat{\Theta}_{ai}^T(t)\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}$$
$$\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}(t)+\sum_{i=1}^{n}\left(\frac{k_i}{4}\|\tilde{x}_i(t)\|^2+\frac{1}{2}\left\|\hat{\Theta}_{fi}^T\varphi_{fi}(x_i)\right\|^2\right.$$
$$\left.+\frac{1}{2}\|\dot{x}_d\|^2+\frac{\kappa_{ai}c_i+2}{2}\left\|\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\Theta_i^*\right\|^2\right). \quad (67)$$

Substituting the facts that

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

into (67) yields

$$\dot{E}(t)\le-\sum_{i=1}^{n}(\gamma_i-k_i-2)\|\hat{e}_i(t)\|^2-\sum_{i=1}^{n}\frac{\kappa_{ai}c_i}{2}\tilde{\Theta}_{ai}^T(t)$$
$$\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\tilde{\Theta}_{ai}(t)-\sum_{i=1}^{n}\frac{\kappa_{ci}}{2\left(1+\|\xi_i\|^2\right)}\tilde{\Theta}_{ci}^T\xi_i\xi_i^T\tilde{\Theta}_{ci}$$
$$+\sum_{i=1}^{n}\frac{c_i\kappa_{ci}}{4\left(1+\|\xi_i\|^2\right)}\tilde{\Theta}_{ai}^T(t)\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\Theta_i^{*T}\xi_i\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}(t)$$
$$+\sum_{i=1}^{n}\frac{c_i\kappa_{ci}}{4\left(1+\|\xi_i\|^2\right)}\tilde{\Theta}_{ci}^T(t)\xi_i\Theta_i^{*T}\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\tilde{\Theta}_{ai}(t)$$
$$-\sum_{i=1}^{n}\frac{\kappa_{ai}c_i}{2}\hat{\Theta}_{ai}^T(t)\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}(t)+\psi_e(t) \quad (68)$$

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:

$$\frac{\kappa_{ci}c_i}{4\left(1+\|\xi_i(t)\|^2\right)}\tilde{\Theta}_{ai}^T(t)\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\Theta_i^{*T}\xi_i(t)\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}(t)$$
$$\le\frac{c_i}{32}\tilde{\Theta}_{ai}^T(t)\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\Theta_i^{*T}\xi_i(t)\xi_i^T(t)\Theta_i^*\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\tilde{\Theta}_{ai}(t)$$
$$+\frac{\kappa_{ci}^2c_i}{2}\hat{\Theta}_{ai}^T(t)\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}(t),$$
$$\frac{\kappa_{ci}c_i}{4\left(1+\|\xi_i(t)\|^2\right)}\tilde{\Theta}_{ci}^T(t)\xi_i(t)\Theta_i^{*T}\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\tilde{\Theta}_{ai}(t)$$
$$\le\frac{c_i}{32\left(1+\|\xi_i(t)\|^2\right)}\tilde{\Theta}_{ci}^T(t)\xi_i(t)\Theta_i^{*T}\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}$$
$$\Theta_i^*\xi_i^T(t)\tilde{\Theta}_{ci}(t)+\frac{\kappa_{ci}^2c_i}{2}\tilde{\Theta}_{ai}^T(t)\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\tilde{\Theta}_{ai}(t).$$

Inserting the above-mentioned facts into (68) yields

$$\dot{E}(t) \le -\sum_{i=1}^{n}(\gamma_i - k_i - 2)\|\hat{e}_i\|^2 - \sum_{i=1}^{n}\left(\frac{\kappa_{ai}c_i}{2} - \frac{\kappa_{ci}^2 c_i}{2}\right.$$
$$\left. - \frac{c_i}{32}\Theta_i^{*T}\xi_i(t)\xi_i^T(t)\Theta_i^*\right)\tilde{\Theta}_{ai}^T(t)\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\tilde{\Theta}_{ai}(t)$$
$$-\sum_{i=1}^{n}\frac{1}{\left(1+\|\xi_i\|^2\right)}\left(\frac{\kappa_{ci}}{2} - \frac{c_i}{32}\Theta_i^{*T}\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\Theta_i^*\right)$$
$$\tilde{\Theta}_{ci}^T(t)\xi_i\xi_i^T\tilde{\Theta}_{ci}(t) - \sum_{i=1}^{n}\left(\frac{\kappa_{ai}c_i}{2} - \frac{\kappa_{ci}^2 c_i}{2}\right)\hat{\Theta}_{ai}^T(t)\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}$$
$$\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\hat{\Theta}_{ai}(t) + \psi_e(t). \tag{69}$$

Make the design parameters to satisfy the following conditions:

$$\gamma_i \ge k_i + 2, \kappa_{ci} \ge \frac{c_i}{16}\Theta_i^{*T}\frac{\partial\varphi_i(\hat{e}_i)}{\partial\hat{e}_i^T}\frac{\partial^T\varphi_i(\hat{e}_i)}{\partial\hat{e}_i}\Theta_i^*,$$
$$\kappa_{ai} \ge \kappa_{ci}^2 + \frac{\zeta_i}{16}\Theta_i^{*T}\Theta_i^*. \tag{70}$$

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

$$\dot{E}(t) \le -\sum_{i=1}^{n}(\gamma_i - k_i - 2)\|\hat{e}_i(t)\|^2 - \sum_{i=1}^{n}\left(\frac{\kappa_{ai}c_i}{2} - \frac{\kappa_{ci}^2 c_i}{2}\right.$$
$$\left. - \frac{\zeta_i c_i}{32}\Theta_i^{*T}\Theta_i^*\right)\lambda_i^{\min}\tilde{\Theta}_{ai}^T(t)\tilde{\Theta}_{ai}(t) - \sum_{i=1}^{n}\left(\frac{\kappa_{ci}\varsigma_i}{2} - \frac{\lambda_i^{\max}\varsigma_i c_i}{32}\right.$$
$$\left.\Theta_i^{*T}\Theta_i^*\right)\tilde{\Theta}_{ci}^T(t)\tilde{\Theta}_{ci}(t) + \psi_e(t) \tag{71}$$

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

$$\dot{E}(t) \le -\gamma\sum_{i=1}^{n}\|\hat{e}_i(t)\|^2 - \kappa_a\sum_{i=1}^{n}\tilde{\Theta}_{ai}^T(t)\tilde{\Theta}_{ai}(t)$$
$$-\kappa_c\sum_{i=1}^{n}\tilde{\Theta}_{ci}^T(t)\tilde{\Theta}_{ci}(t) + \beta_e. \tag{72}$$

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

$$\dot{E}(t) \le -\frac{\gamma}{\lambda_{\max}}\hat{z}^T(t)(\tilde{L}\otimes I_m)\hat{z}(t) - \kappa_a\sum_{i=1}^{n}\tilde{\Theta}_{ai}^T(t)\tilde{\Theta}_{ai}(t)$$
$$-\kappa_c\sum_{i=1}^{n}\tilde{\Theta}_{ci}^T(t)\tilde{\Theta}_{ci}(t) + \beta_e \le -\alpha_e E(t) + \beta_e \tag{73}$$

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

According to Lemma 4, there is the fact that

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

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

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

to (75) has

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

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 $\psi_z(t) \le \beta_z$. Furthermore, based on (80) (in Remark 1), there is the following equation:

$$\dot{E}_z(t) \le -\frac{\gamma}{\lambda_{\max}}\hat{z}^T(t)(\tilde{L}\otimes I_m)\hat{z}(t) + \beta_z$$
$$= -\alpha_z E_z(t) + \beta_z \tag{77}$$

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

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

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

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. □

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 $\tilde{L} = Q^T PQ$. Then the term $\hat{z}^T(t)(\tilde{L} \otimes I_m)\hat{z}(t)$ can be reexpressed as

$$\begin{aligned}\hat{z}^T(t)(\tilde{L} \otimes I_m)\hat{z}(t) &= \hat{z}^T(t)\left((Q^T PQ) \otimes I_m\right)\hat{z}(t)\\ &= \hat{z}^T(t)\left((Q^T PQQ^T P^{-1}QQ^T PQ) \otimes I_m\right)\hat{z}(t)\\ &= \hat{z}^T(t)(\tilde{L} \otimes I_m)^T\left((Q^T P^{-1}Q) \otimes I_m\right)(\tilde{L} \otimes I_m)\hat{z}(t)\\ &= \hat{e}^T(t)\left((Q^T P^{-1}Q) \otimes I_m\right)\hat{e}(t).\end{aligned} \tag{79}$$

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

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

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:

$$\dot{x}_i(t) = -\alpha_i x_i(t) - \begin{bmatrix} 0.5x_{i1}\cos^2(\beta_i x_{i1}) \\ x_{i2} - \sin^2(\beta_i x_{i2}) \end{bmatrix} + u_i,\quad i = 1, 2, 3, 4 \tag{81}$$

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

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

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

$$\mu^i_{F^j}(x_i) = \exp\left(-\frac{\left\|x_i - [6, 6]^T + [2j - 1, 2j - 1]^T\right\|^2}{2}\right),\quad j = 1, \ldots, 6. \tag{83}$$

Then the fuzzy basis function vector is obtained as $\varphi_{fi}(x_i) = [\varphi^1_{fi}(x_i), \ldots, \varphi^6_{fi}(x_i)]$, where $\varphi^j_{fi}(x_i) = \frac{\mu^i_{F^j}(x_i)}{\sum_{j=1}^6 \mu^i_{F^j}(x_i)}$, $j = 1, \ldots, 6$. Based on (23), the adaptive identifier is built in the following by choosing the design parameters $k_{i=1,2,3,4} = 24, 20, 18, 16$; $\Gamma_{i=1,2,3,4} = 0.4I_6$; and $\sigma_{i=1,2,3,4} = 0.6$:

$$\begin{aligned}\dot{\hat{x}}_i(t) &= -k_i\tilde{x}_i(t) + \hat{\Theta}^T_{fi}(t)\varphi_{fi}(x_i) + u_i,\\ \dot{\hat{\Theta}}_{fi}(t) &= 0.4\left(-\varphi_{fi}(x_i)\tilde{x}^T_i(t) - 0.6\hat{\Theta}_{fi}(t)\right)\end{aligned} \tag{84}$$

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

$$\mu^i_{F^j}(e_i) = \exp\left(-\frac{\left\|e_i - [6, 6]^T + [2j - 1, 2j - 1]^T\right\|^2}{2}\right),\quad j = 1, \ldots, 6. \tag{85}$$

The fuzzy basis function vector is yielded as $\varphi_i(e_i) = [\varphi^1_i(e_i), \ldots, \varphi^6_i(e_i)]$, where $\varphi^j_i(e_i) = \frac{\mu^i_{F^j}(e_i)}{\sum_{j=1}^6 \mu^i_{F^j}(e_i)}$. 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:

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

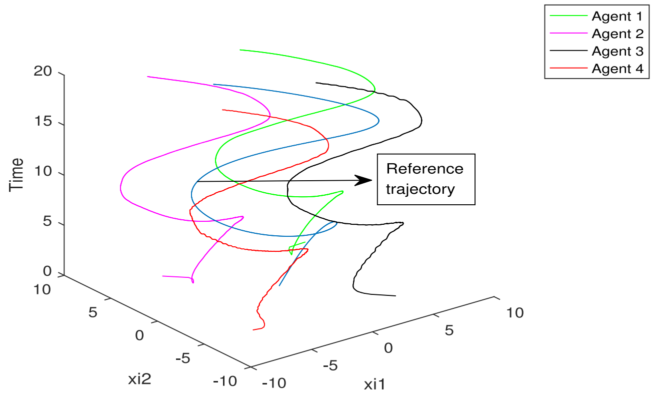


Fig. 1. Multi-agent formation performance.

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.

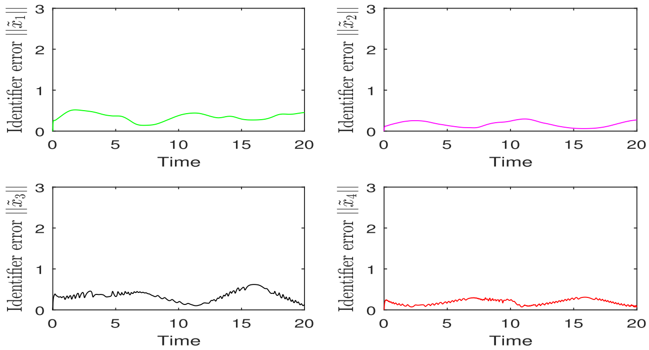


Fig. 2. Norm $\|\tilde{x}_i\|$, $i = 1, 2, 3, 4$, of the adaptive identifier error.

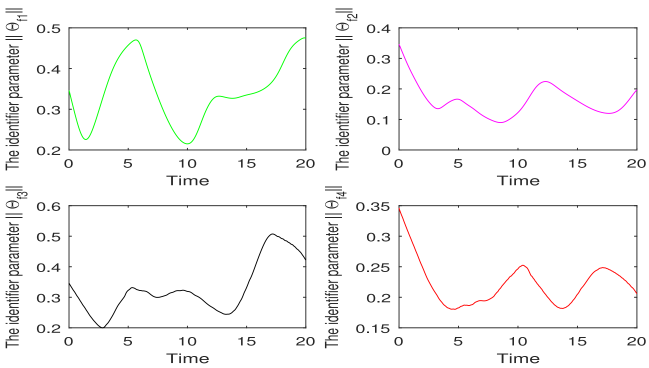


Fig. 3. Norm $\left\|\hat{\Theta}_{fi}\right\|$, $i = 1, 2, 3, 4$, of the identifier parameter matrix.

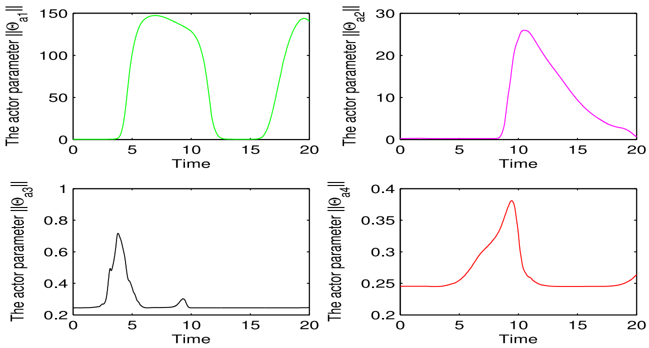


Fig. 4. Norm $\left\|\hat{\Theta}_{ai}\right\|$, $i = 1, 2, 3, 4$, of the actor parameter vector.

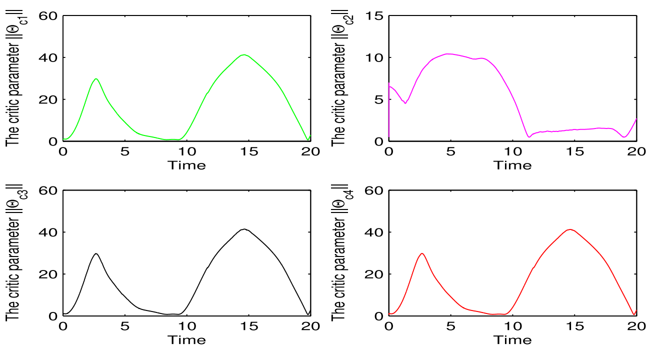


Fig. 5. Norm $\left\|\hat{\Theta}_{ci}\right\|$, $i = 1, 2, 3, 4$, of the critic parameter vector.

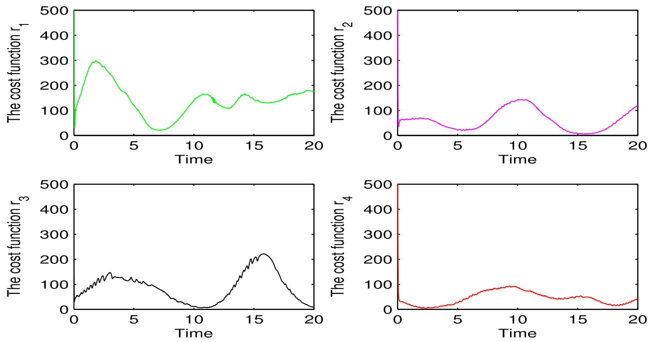


Fig. 6. Cost function $r_i(x_i, u_i)$, $i = 1, 2, 3, 4$.

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.

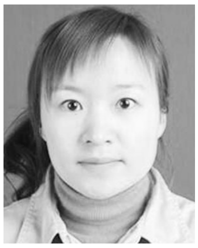
[34] G. H. Hardy, J. E. Littlewood, and G. Plya, *Inequalities (Cambridge Mathematical Library)*. Cambridge, U.K.: Cambridge Univ. Press, 1934.

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.

报告编号：BZU20180031

文献检索证明

作者姓名：文国兴

作者单位：滨州学院

该作者 2018 年发表在期刊《IEEE TRANSACTIONS ON FUZZY SYSTEMS》第 26 卷第 5 期的论文《Optimized Multi-Agent Formation Control Based on an Identifier-Actor--Critic Reinforcement Learning Algorithm》被 SCIE（SCI-EXPANDED）收录。该期刊 2017 年 JCR 的影响因子是 8.415，JCR 分区见附件。在中科院 2016 年 JCR 分区表中位于大类分区 1 区。

检索结果见附件。

证明单位：滨州学院图书馆

证明人：　郭春霞

2018 年 10 月 26 日

关闭

Web of Science
第 1 页 (记录 1 -- 1)
◀[1]▶

打印

第 1 条，共 1 条

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

作者: Wen, GX (Wen, Guoxing); Chen, CLP (Chen, C. L. Philip); Feng, J (Feng, Jun); Zhou, N (Zhou, Ning)

来源出版物: IEEE TRANSACTIONS ON FUZZY SYSTEMS 卷: 26 期: 5 页: 2719-2731 **DOI:** 10.1109/TFUZZ.2017.2787561 出版年: OCT 2018

Web of Science 核心合集中的 "被引频次": 0

被引频次合计: 0

入藏号: WOS:000446675400019

文献类型: Article

地址: [Wen, Guoxing] Binzhou Univ, Coll Sci, Binzhou 256600, Peoples R China.
[Chen, C. L. Philip] Univ Macau, Fac Sci & Technol, Dept Comp & Informat Sci, Macau 99999, Peoples R China.
[Chen, C. L. Philip] Dalian Maritime Univ, Dalian 116026, Peoples R China.
[Chen, C. L. Philip] Chinese Acad Sci, Inst Automat, State Key Lab Management & Control Complex Syst, Beijing 100080, Peoples R China.
[Feng, Jun] Nanjing Univ Aeronaut & Astronaut, Coll Comp Sci & Technol, Nanjing 210000, Jiangsu, Peoples R China.
[Feng, Jun] Binzhou Univ, Dept Informat Engn, Binzhou 256600, Peoples R China.
[Zhou, Ning] Univ Groningen, Fac Sci & Engn, NL-9747 AG Groningen, Netherlands.
[Zhou, Ning] Fujian Agr & Forestry Univ, Coll Comp & Informat Sci, Fuzhou 350002, Fujian, Peoples R China.

通讯作者地址: Wen, GX (通讯作者)， Binzhou Univ, Coll Sci, Binzhou 256600, Peoples R China.

电子邮件地址: gxwen@live.cn; philip.chen@ieee.org; fengjunImportant@163.com; zhouning2010@gmail.com

ISSN: 1063-6706

eISSN: 1941-0034

来源出版物页码计数: 13

关闭

Web of Science
第 1 页 (记录 1 -- 1)
◀[1]▶

打印

IEEE TRANSACTIONS ON FUZZY SYSTEMS

impact factor

8.415 9.34
2017 5 年

JCR®类别	类别中的排序	JCR分区
COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE	4/132	Q1
ENGINEERING, ELECTRICAL & ELECTRONIC	7/260	Q1

数据来自第 2017 版 Journal Citation Reports

出版商

IEEE-INST ELECTRICAL ELECTRONICS ENGINEERS INC, 445 HOES LANE, PISCATAWAY, NJ 08855-4141 USA

ISSN: 1063-6706
eISSN: 1941-0034

研究领域

Computer Science
Engineering

关闭窗口

Optimized Backstepping for Tracking Control of Strict-Feedback Systems

Guoxing Wen, Shuzhi Sam Ge, *Fellow, IEEE*, and Fangwen Tu

***Abstract*— In this paper, a control technique named optimized backstepping is first proposed by implementing tracking control for a class of strict-feedback systems, which considers optimization as a design philosophy of the high-order system control. The basic idea is that designing the actual and virtual controls of backstepping is the optimized solutions of the corresponding subsystems so that overall control of the high-order system is optimized. In general, optimization control is designed based on the solution of Hamilton–Jacobi–Bellman equation, but solving the equation is very difficult due to the inherent nonlinearity and intractability. In order to overcome the difficulty, the neural network (NN)-based reinforcement learning strategy of actor–critic architecture is used. In every backstepping step, the actor and critic NNs are constructed for executing control behavior and evaluating control performance, respectively. According to the Lyapunov stability theorem, it is proven that the desired control performance can be obtained. Finally, a simulation example is carried out to further demonstrate the effectiveness of the proposed control approach.**

***Index Terms*— Actor–critic architecture, Lyapunov stability, optimized backstepping (OB), strict-feedback system, tracking control.**

I. Introduction

AFTER decades of research and development, backstepping has become the most common and powerful control strategy for strictly feedback and lower triangular systems, and it also provides a systematic theory framework for the practical engineering [1]–[4]. Its basic idea is to construct a recursive control by considering many state variables as "virtual control" and designing the control laws for them so that the goals of stabilizing and tracking are achieved by an ordered control sequence, and then, it is proven by performing the Lyapunov stability analysis for the entire system.

In the past decade, many representative results concerning backstepping control have been reported, such as [5]–[9]. In [5], a multilayer neural network (NN)-based adaptive control for nonlinear strict-feedback systems is addressed by using backstepping technique. In [6], an adaptive control scheme for nonlinear strict-feedback systems is developed by combining both dynamic surface and backstepping techniques. In [7], backstepping technique is applied to consensus tracking control of high-order nonlinear multiagent systems. In [8] and [9], the cooperative control of high-order nonlinear stochastic multiagent systems is investigated, and this is a very challenging work, because the exogenous disturbances depicted by Wiener process are considered. Finally, it is proven that the control objective can be accomplished by applying backstepping techniques. Although the backstepping technique had been well developed and applied in control community, unfortunately, all backstepping-based controls never address the optimization problem so far. Motivated by the above-mentioned considerations, a high-order system control technique named optimized backstepping (OB) is first proposed by implementing tracking control for a class of strict-feedback systems. Since the actual and virtual controls are designed to be the optimized solutions of corresponding subsystems, the overall control is optimized.

Ever since optimal control was formally developed about five decades ago by Bellman [10] and Pontryagin [11], optimization has become a fundamental principles and gained increasing attention in modern control theory. Optimal control means that the cost function is minimized by a control protocol. In general, an optimal controller is designed based on a gradient of the optimal value function, which is expected to obtain by solving Hamilton–Jacobi–Bellman (HJB) equation [12], and becomes Riccati equation for the linear system. However, HJB equation is solved difficultly by analytical approaches owing to the inherent nonlinearity and intractability.

In the last decades, reinforcement learning (RL)-based function approximation strategy [13] has been successfully applied to adaptive optimization control and becomes a popular approach to solve the complex control problem. In brief, RL is that the appropriate actions are obtained by evaluating the feedback from the environment. One of the well-known and effective means is actor–critic architecture, where the actor performs certain actions by interacting with environment; the critic evaluates the actions and returns feedback to the actor so that the performance of subsequent actions is improved [14]. Actor–critic RL as one of the most powerful and popular online learning approaches has been widely applied to optimization control, such as [15]–[17] for continuous systems and [18] and [19] for discrete systems.

Manuscript received June 24, 2017; revised December 20, 2017; accepted January 30, 2018. Date of publication March 6, 2018; date of current version July 18, 2018. This work was supported in part by the Doctoral Scientific Research Staring Fund of Binzhou University under Grant 2016Y14 and in part by the National Natural Science Foundation of China under Grant 61703050. *(Corresponding author: Guoxing Wen.)*

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

S. S. Ge is with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore 117576, and also with the Social Robotics Laboratory, Interactive Digital Media Institute, National University of Singapore, Singapore 117576 (e-mail: samge@nus.edu.sg).

F. Tu is with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore 117576 (e-mail: fangwen_tu@hotmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNNLS.2018.2803726

Based on RL, approximate dynamic programming (ADP), which was first developed by Werbos [20], was successfully applied to adaptive optimization control by using optimal value function approximation (typically NN-based approximation). In recent years, iterative ADP optimal control methods are well investigated by using actor–critic-based approximation methods, and several excellent research results have attracted considerable attention [21]–[23]. The research in [21] concerns with developing an online approximate solution for continuous-time nonlinear systems by actor–critic NNs, where both actor and critic NNs are trained simultaneously. In [22], an integral RL algorithm of the actor–critic structure is developed to solve HJB equation for partially unknown constrained-input systems. In [23], the problem of system parameter uncertainties has been tackled by using an iterative actor–critic method.

These optimization schemes have attracted considerable attention from different research fields and have been widely applied in practical engineering. However, the optimization control of high-order systems is still rarely addressed, especially for tracking control because of the difficulties coming from controller design and performance analysis. Motivated by the above-mentioned discussions, an optimizing control technique of high-order systems is proposed by performing tracking control. Applying the universal approximation ability of NNs, both the actor and critic NNs are constructed to carry out the RL algorithm, in which actor NN is trained for obtaining excellent system stability performance and critic NN is tuned based on minimizing Bellman error. The main contributions are listed in the following.

1) The proposed OB control technique can achieve the optimized control of high-order systems by melting optimization into backstepping control. The basic idea is that every controller is designed to be the optimized solution of corresponding subsystem, and therefore, the overall system control is optimized.
2) The proposed optimized approach can efficiently solve tracking control problem by segmenting an error term from the optimal value function. Owing to the difficulty coming from the convergence analysis of tracking errors, existing optimization control methods rarely involve the tracking problem. By both theory proof and computer simulation, it is demonstrated that the control scheme can steer the system output to follow the desired trajectory.
3) Online RL is applied to backstepping control. By evaluating the feedback from environment and returning the evaluation to facilitate the control, excellent control performance can be guaranteed.

II. Preliminaries

A. Background on Optimal Control

Consider the nonlinear continuous-time dynamic system modeled in the following:

$$\dot{x}(t) = f(x) + g(x)u(x) \tag{1}$$

where $x(t) \in R^n$ is the state vector, $u(x) \in R^m$ is the control input, and $f(x) \in R^n$ with $f(0_n) = 0_n$ and $g(x) \in R^{n\times m}$ are the vector-valued and matrix-valued functions, respectively. The term $f(x) + g(x)u(x)$ is assumed Lipschitz continuous on the set Ω containing origin so that the solution of (1) is unique for bounded initial state $x(0)$ and control input $u(x)$. The system (1) is required stabilizable on Ω, i.e., there exists the continuous control function $u(x)$ such that the system is asymptotically stable.

The infinite horizon value function for the dynamic system (1) is defined as

$$V(x) = \int_t^\infty r(x(s), u(x))ds \tag{2}$$

where $r(x(t), u(x)) = h(x) + u^T Pu$ is the immediate or local cost function, of which $h(x)$ is a positive definite function and $h(x) = 0$ if and only if $x(t) = 0$, and $P \in R^{m\times m}$ is a positive definite matrix.

Definition 1 [24]: A control policy $u(x)$ is defined as admissible with respect to (1) on Ω, which is denoted by $u(x) \in \Psi(\Omega)$, if $u(x)$ is continuous on Ω with $u(0) = 0$, $u(x)$ stabilizes (1) on Ω, and $V(x)$ is finite.

The optimal control problem for system (1) is to find a control policy $u(x) \in \Psi(\Omega)$, such that the infinite horizon value function (2) is minimized.

Define the Hamiltonian function corresponding to (1) and (2) as

$$\begin{aligned} H(x, u, V_x) &= r(x, u) + V_x^T(x)\dot{x}(t) \\ &= h(x) + u^T(x)Pu(x) + V_x^T(x) \\ &\quad \times (f(x) + g(x)u(x)) \end{aligned}$$

where $V_x(x) = \partial V(x)/\partial x$, i.e., the gradient of function $V(x)$ with respect to $x(t)$.

The optimal value function is defined as

$$\begin{aligned} V^*(x) &= \min_{u\in\Psi(\Omega)} \left(\int_t^\infty r(x(s), u(x))ds \right) \\ &= \int_t^\infty r(x(s), u^*(x))ds \end{aligned}$$

where $u^*(x)$ is the optimal controller. Then, there is the following HJB equation:

$$\begin{aligned} &H\left(x, u^*, V_x^*\right) \\ &\quad = r(x, u^*) + V_x^{*T}(x)\dot{x}(t) \\ &\quad = h(x) + u^{*T} Pu^* + V_x^{*T}(x)(f(x) + g(x)u^*) = 0 \end{aligned} \tag{3}$$

where $V_x^*(x) = \partial V^*(x)/\partial x$, i.e., the gradient of $V^*(x)$ with respect to x.

Assuming the solution of (3) existent and unique, the optimal control u^* can be obtained by solving $\partial H(x, u^*, V_x^*)/\partial u^* = 0$ as

$$u^*(x) = -\frac{1}{2}P^{-1}g^T(x)V_x^*(x). \tag{4}$$

Substituting (4) into (3), the following result can be obtained:

$$\begin{aligned} H\left(x, u^*, V_x^*\right) &= h(x) + V_x^{*T} f(x) \\ &\quad - \frac{1}{4}V_x^{*T}(x)g(x)P^{-1}g^T(x)V_x^*(x) = 0. \end{aligned} \tag{5}$$

The gradient term $V_x^*(x)$ is expected to obtain by solving (5), and then, the optimal control can be got by substituting the solution into (4). However, it is very difficult or impossible to solve the HJB equation (5) owing to the inherent nonlinearity and intricacy. In order to overcome the difficulty of solving HJB equation, the adaptive approximation strategy using RL of actor–critic architecture is usually considered [25].

B. Neural Networks and Function Approximation

It has been proven that NNs have excellent function approximation and adaptive learning abilities. Any continuous nonlinear function $\varphi(z) : R^n \to R^m$ defined on a compact set Ω_z can be approximated by NNs in the following form:

$$\varphi_{NN}(z) = W^T S(z) \tag{6}$$

where $W \in R^{p\times m}$ is the weight matrix and p is the neuron number; $S(z) = [s_1(z), \ldots, s_p(z)]^T$ is the basis function vector, $s_i(z) = \exp[-(z-\mu_i)^T(z-\mu_i)/\phi_i^2]$, $z \in \Omega_z \subset R^n$, is the input vector, ϕ_i is the width of Gaussian function, and $\mu_i = [\mu_{i1}, \ldots, \mu_{in}]^T$, μ_{ij} is the center of receptive field, $i = 1, 2, \ldots, p$, $j = 1, 2, \ldots, n$.

Based on the NN approximation (6), the function $\varphi(z)$ can be redescribed in the following form:

$$\varphi(z) = W^{*T} S(z) + \varepsilon(z) \tag{7}$$

where $\varepsilon(z) \in R^m$ is the approximation error, which is bounded by a positive constant δ, i.e., $\|\varepsilon(z)\| \le \delta$; $W^* \in R^{p\times m}$ is the ideal weight, which is defined as $W^* \triangleq \arg\min_{W\in R^{p\times m}}\{\sup_{z\in\Omega_z}\|\varphi(z) - WS(z)\|\}$. It should be mentioned that the ideal NN weight W^* is an "artificial" quantity only for analysis purpose.

It has been demonstrated that the approximation error $\|\varepsilon(z)\|$ can be reduced to arbitrarily small if the neuron number p is chosen large enough [26].

III. Main Result

A. Problem Description

Consider the following single-input single-output nonlinear strict-feedback system:

$$\begin{aligned}\dot{x}_1(t) &= f_1(\bar{x}_1(t)) + x_2(t)\\ \dot{x}_2(t) &= f_2(\bar{x}_2(t)) + x_3(t)\\ &\cdots\\ \dot{x}_n(t) &= f_n(\bar{x}_n(t)) + u\end{aligned} \tag{8}$$

where $x_1 \in R$ is the system output, $u \in R$ is the control input, $\bar{x}_i(t) = [x_1(t), \ldots, x_i(t)]^T \in R^i$ is the state vector, $f_i(\bar{x}_i) \in R$ with $f_i(\bar{0}_i) = 0$ is the continuous function, which is assumed known and bounded, and $f_i(\bar{x}_i)+x_{i+1}(t)$, $i = 1, \ldots, n-1$, and $f_n(\bar{x}_n) + u$ are assumed Lipschitz continuous and stabilizable on the sets containing origin.

Definition 2 (Semi-Globally Uniformly Ultimately Bounded [27]): Consider the nonlinear system

$$\dot{x}(t) = f(x, t)$$

where $x(t) \in R^n$ is the state vector. Its solution is said to be semi-globally uniformly ultimately bounded (SGUUB) if, for $x(0) \in \Omega_x$ where $\Omega_x \in R^n$ is a compact set, there exist two constants σ and $T(\sigma, x(0))$, such that $\|x(t)\| \le \sigma$ is held for all $t > t_0 + T(\sigma, x(0))$.

Lemma 1 [28]: Let $G(t) \in R$ be a continuous positive function with bounded initial value $G(0)$. If $\dot{G}(t) \le -aG(t)+c$ is held, where a and c are two positive constants, then there is the following one:

$$G(t) \le e^{-at}G(0) + \frac{c}{a}(1 - e^{-at}).$$

The control objective is to design the NN approximation-based optimized control for the strict-feedback system (8), such that: 1) all error signals of the tracking control are SGUUB and 2) the system output $x_1(t)$ can track the reference signal $y_r(t)$ to the desired accuracy, where $y_r(t)$ is a sufficiently smooth function and $y_r(t), \dot{y}_r(t), \cdots, y_r^{n-1}(t)$ are bounded.

B. Optimized Backstepping Design

In this section, optimizing is integrated into the n-step backstepping for the tracking control of the strict-feedback system (8). Different with traditional backstepping, the proposed OB control designs all virtual controls and the actual control to be the optimized solutions of corresponding subsystems, therefore the overall system control can be optimized. In every backstepping step, the actor–critic RL algorithm is implemented by constructing both actor and critic NNs, where the actor NN is utilized to perform the control policy and the critic NN is utilized to evaluate the optimization performance.

Step 1: Define the tracking error variable for the backstepping step as $z_1(t) = x_1(t) - y_r(t)$. Its time derivative along (8) is

$$\dot{z}_1(t) = f_1(\bar{x}_1(t)) + x_2(t) - \dot{y}_r(t) \tag{9}$$

where $x_2(t)$ is called the intermediate controller.

Viewing $x_2(t)$ as the optimal virtual control $\alpha_1^*(z_1)$, i.e., $x_2(t) \triangleq \alpha_1^*(z_1)$, the optimal value function for z_1-subsystem (9) is defined as

$$\begin{aligned}V_1^*(z_1) &= \min_{\alpha_1\in\Psi(\Omega_{z_1})}\left(\int_t^\infty r_1(z_1(s), \alpha_1(z_1))ds\right)\\ &= \int_t^\infty r_1\big(z_1(s), \alpha_1^*(z_1)\big)ds\end{aligned} \tag{10}$$

where $r_1(z_1, \alpha_1) = z_1^2(t) + \alpha_1^2(z_1)$ is the cost function, $\alpha_1(z_1)$ is the virtual controller, and Ω_{z_1} is a compact set containing origin. The optimal value function can be decomposed into the following two terms:

$$\begin{aligned}V_1^*(z_1) &= \beta_1 z_1^2(t) - \beta_1 z_1^2(t) + V_1^*(z_1)\\ &= \beta_1 z_1^2(t) + V_1^o(z_1)\end{aligned} \tag{11}$$

where β_1 is the positive design constant and $V_1^o(z_1) \in R$ is a continuous scalar-value function.

Remark 1: The decomposed term $\beta_1 z_1^2(t)$of (11), which will be made in every step later, aims to achieve the tracking control for the subsystem. Although most existing optimization control methods, such as [15]–[17], can guarantee state boundedness and system stability, few research results address tracking control problems. In the design, by segmenting the error term $\beta_1 z_1^2(t)$ from the optimal value function and choosing the appropriate parameter β_1, the desired tracking performance can be achieved. The method can also be easily extended to multidimensional systems by changing the term $\beta_1 z_1^2(t)$ to norm expression.

Based on the error dynamic (9), HJB equation of the z_1-subsystem is

$$\begin{aligned} H_1\left(z_1, \alpha_1^*, \frac{\partial V_1^*}{\partial z_1}\right) &= r_1(z_1, \alpha_1^*) + \frac{\partial V_1^*(z_1)}{\partial z_1}\dot{z}(t) \\ &= z_1^2(t) + \alpha_1^{*2}(z_1) + \left(2\beta_1 z_1(t) + \frac{\partial V_1^o(z_1)}{\partial z_1}\right) \\ &\quad \times \left(f_1(\bar{x}_1) + \alpha_1^*(z_1) - \dot{y}_r(t)\right) = 0. \end{aligned} \tag{12}$$

By solving $\partial H_1/\partial \alpha_1^* = 0$, the optimal control α_1^* is

$$\alpha_1^*(z_1) = -\beta_1 z_1(t) - \frac{1}{2}\frac{\partial V_1^o(z_1)}{\partial z_1}. \tag{13}$$

Since the term $\partial V_1^o(z_1)/\partial z_1$ is continuous on the compact set Ω_{z_1}, it can be approximated by NN as

$$\frac{\partial V_1^o(z_1)}{\partial z_1} = W_1^{*T} S_1(z_1) + \varepsilon_1(z_1) \tag{14}$$

where $W_1^* \in R^{m_1}$ is the ideal weight, $S_1(z_1) \in R^{m_1}$ is the basis function vector, and $\varepsilon_1(z_1) \in R$ is the approximation error, which is bounded by a constant δ_1, i.e., $|\varepsilon_1(z_1)| \le \delta_1$.

Using (14), the gradient term $\partial V_1^*(z_1)/\partial z_1$ and the optimal controller $\alpha_1^*(z_1)$ become

$$\frac{\partial V_1^*(z_1)}{\partial z_1} = 2\beta_1 z_1(t) + W_1^{*T} S_1(z_1) + \varepsilon_1(z_1) \tag{15}$$

$$\alpha_1^*(z_1) = -\beta_1 z_1(t) - \frac{1}{2}\left(W_1^{*T} S_1(z_1) + \varepsilon_1(z_1)\right). \tag{16}$$

Substituting (14) and (16) into (12), the following equation can be obtained:

$$\begin{aligned} H_1&\left(z_1, \alpha_1^*, W_1^*\right) \\ &= -\left(\beta_1^2 - 1\right) z_1^2(t) + 2\beta_1 z_1(t)(f_1(\bar{x}_1) - \dot{y}_r(t)) \\ &\quad + W_1^{*T} S_1(z_1)(f_1(\bar{x}_1(t)) - \dot{y}_r(t) - \beta_1 z_1(t)) \\ &\quad - \frac{1}{4} W_1^{*T} S_1(z_1) S_1^T(z_1) W_1^* + \epsilon_1(t) = 0 \end{aligned} \tag{17}$$

where $\epsilon_1(t) = \varepsilon_1(z_1)(f_1(\bar{x}_1) - \dot{y}_r(t) + \alpha_1^*) + (1/4)\varepsilon_1^2(z_1)$ is bounded, because all its terms are bounded.

Since the ideal weight W_1^* is an unknown constant vector, the optimal virtual control (16) cannot be applied directly. In order to obtain the available control, the actor–critic RL algorithm is constructed, where the critic and actor NNs are given in the following:

$$\frac{\partial \hat{V}_1^*}{\partial z_1} = 2\beta_1 z_1(t) + \frac{\partial \hat{V}_1^o}{\partial z_1} = 2\beta_1 z_1(t) + \hat{W}_{c1}^T(t) S_1(z_1) \tag{18}$$

$$\hat{\alpha}_1(z_1) = -\beta_1 z_1(t) - \frac{1}{2}\hat{W}_{a1}^T(t) S_1(z_1) \tag{19}$$

where $\hat{V}_1^*(z_1)$ and $\hat{V}_1^o(z_1)$ are the estimations of $V_1^*(z_1)$ and $V_1^o(z_1)$, respectively; $\hat{W}_{c1}^T(t) \in R^{m_1}$ and $\hat{W}_{a1}^T(t) \in R^{m_1}$ are the critic and actor NN weights, respectively.

By substituting (18) and (19) into (12), the approximated HJB equation can be obtained in the following:

$$\begin{aligned} H_1&(z_1, \hat{\alpha}_1, \hat{W}_{c1}) \\ &= z_1^2(t) + \left(-\beta_1 z_1(t) - \frac{1}{2}\hat{W}_{a1}^T S_1(z_1)\right)^2 \\ &\quad + \left(2\beta_1 z_1(t) + \hat{W}_{c1}^T S_1(z_1)\right)\left(-\beta_1 z_1(t) - \frac{1}{2}\hat{W}_{a1}^T(t) S_1(z_1)\right. \\ &\quad \left. + f_1(\bar{x}_1) - \dot{y}_r(t)\right). \end{aligned} \tag{20}$$

Using the HJB equation (17) and its approximation (20), Bellman residual error $e_1(t)$ is yielded as

$$\begin{aligned} e_1(t) &= H_1(z_1, \hat{\alpha}_1, \hat{W}_{c1}) - H_1\left(z_1, \alpha_1^*, W_1^*\right) \\ &= H_1(z_1, \hat{\alpha}_1, \hat{W}_{c1}). \end{aligned} \tag{21}$$

Define a positive definite function of the error $e_1(t)$ as

$$E_1(t) = \frac{1}{2}e_1^2(t). \tag{22}$$

In order to minimize the Bellman error (21), the following critic NN updating law is obtained by using the gradient descent algorithm:

$$\begin{aligned} \dot{\hat{W}}_{c1}(t) &= -\frac{\gamma_{c1}}{\|\omega_1(t)\|^2 + 1}\frac{\partial E_1(t)}{\partial \hat{W}_{c1}} \\ &= -\frac{\gamma_{c1}}{\|\omega_1\|^2 + 1}\omega_1(t) \\ &\quad \times \left(\omega_1^T(t)\hat{W}_{c1}(t) - \left(\beta_1^2 - 1\right) z_1^2(t) + 2\beta_1 z_1\right. \\ &\quad \left. \times (f_1(\bar{x}_1) - \dot{y}_r) + \frac{1}{4}\hat{W}_{a1}^T S_1(z_1) S_1^T(z_1)\hat{W}_{a1}\right) \end{aligned} \tag{23}$$

where $\gamma_{c1} > 0$ is the learning rate and $\omega_1(t) = S_1(z_1)(f_1(\bar{x}_1) - \beta_1 z_1(t) - (1/2)\hat{W}_{a1}^T(t) S_1(z_1) - \dot{y}_r) \in R^{m_1}$.

Based on the system stability analysis, the actor NN weight updating law is designed as follows:

$$\begin{aligned} \dot{\hat{W}}_{a1}(t) &= \frac{1}{2}S_1(z_1) z_1(t) - \gamma_{a1} S_1(z_1) S_1^T(z_1)\hat{W}_{a1}(t) \\ &\quad + \frac{\gamma_{c1}}{4(\|\omega_1\|^2 + 1)} S_1(z_1) S_1^T(z_1)\hat{W}_{a1}(t)\omega_1^T(t)\hat{W}_{c1}(t) \end{aligned} \tag{24}$$

where $\gamma_{a1} > 0$ is the learning rate.

For obtaining the desired control performance, the following assumption is required.

Assumption 1 (Persistence of Excitation [21]): The signs of $\omega_i(t)\omega_i^T(t)$, $i = 1, 2, \ldots, n$, satisfy the following persistence of excitation (PE) condition over the interval $[t, t + \bar{t}_i]$ with all t values:

$$\eta_i I_{m_i} \le \omega_i(t)\omega_i^T(t) \le \zeta_i I_{m_i} \tag{25}$$

where $\eta_i > 0$, $\zeta_i > 0$, $\bar{t}_i > 0$, and $I_{m_i} \in R^{m_i \times m_i}$ is the identity matrix.

Defining the error variable $z_2(t) = x_2(t) - \hat{\alpha}_1(z_1)$, the error dynamic (9) can be rewritten as

$$\dot{z}_1(t) = f_1(\bar{x}_1) + z_2(t) + \hat{\alpha}_1(z_1) - \dot{y}_r(t). \tag{26}$$

Consider the following Lyapunov function candidate for the z_1-subsystem:

$$L_1(t) = \frac{1}{2}z_1^2(t) + \frac{1}{2}\tilde{W}_{a1}^T(t)\tilde{W}_{a1}(t) + \frac{1}{2}\tilde{W}_{c1}^T(t)\tilde{W}_{c1}(t)$$

where $\tilde{W}_{c1}(t) = \hat{W}_{c1}(t) - W_1^*$ and $\tilde{W}_{a1}(t) = \hat{W}_{a1}(t) - W_1^*$ are the critic and actor NN estimation errors.

The time derivative of $L_1(t)$ along (23), (24), and (26) is

$$\begin{aligned}\dot{L}_1(t) =& z_1(t)(f_1(\bar{x}_1) + z_2(t) + \hat{\alpha}_1(z_1) - \dot{y}_r(t)) + \tilde{W}_{a1}^T(t)\\ &\times\bigg(\frac{1}{2}S_1(z_1)z_1(t) - \gamma_{a1}S_1(z_1)S_1^T(z_1)\hat{W}_{a1}(t)\\ &+\frac{\gamma_{c1}}{4(\|\omega_1\|^2+1)}S_1(z_1)S_1^T(z_1)\hat{W}_{a1}(t)\omega_1^T(t)\hat{W}_{c1}(t)\bigg)\\ &-\frac{\gamma_{c1}}{\|\omega_1\|^2+1}\tilde{W}_{c1}^T(t)\omega_1\\ &\times\bigg(\omega_1^T\hat{W}_{c1}(t) - (\beta_1^2-1)z_1^2(t)\\ &+2\beta_1 z_1(t)(f_1(\bar{x}_1) - \dot{y}_r(t)) + \frac{1}{4}\hat{W}_{a1}^T(t)S_1(z_1)\\ &\times S_1^T(z_1)\hat{W}_{a1}(t)\bigg).\end{aligned} \tag{27}$$

Executing the optimized controller (19) to (27) yields

$$\begin{aligned}\dot{L}_1(t) =& z_1(t)z_2(t) - \beta_1 z_1^2(t) - \frac{1}{2}z_1(t)\hat{W}_{a1}^T(t)S_1(z_1)\\ &+\frac{1}{2}\tilde{W}_{a1}^T(t)S_1(z_1)z_1 - \gamma_{a1}\tilde{W}_{a1}^T(t)S_1(z_1)S_1^T(z_1)\hat{W}_{a1}(t)\\ &+\frac{\gamma_{c1}}{4(\|\omega_1\|^2+1)}\tilde{W}_{a1}^T(t)S_1(z_1)S_1^T(z_1)\hat{W}_{a1}(t)\omega_1^T\hat{W}_{c1}(t)\\ &-\frac{\gamma_{c1}}{\|\omega_1\|^2+1}\tilde{W}_{c1}^T(t)\omega_1\\ &\times\bigg(\omega_1^T\hat{W}_{c1}(t) - (\beta_1^2-1)z_1^2(t)\\ &+2\beta_1 z_1(t)(f_1(\bar{x}_1) - \dot{y}_r) + \frac{1}{4}\hat{W}_{a1}^T(t)S_1(z_1)S_1^T(z_1)\\ &\times\hat{W}_{a1}(t)\bigg) + z_1(t)f_1(\bar{x}_1) - z_1(t)\dot{y}_r.\end{aligned} \tag{28}$$

Based on the fact $\tilde{W}_{a1}(t) = \hat{W}_{a1}(t) - W_1^*$, there are the following equations:

$$\begin{aligned}&\tilde{W}_{a1}^T(t)S_1(z_1)z_1 - z_1\hat{W}_{a1}^T(t)S_1(z_1)\\ &\quad= -z_1(t)W_1^{*T}S_1(z_1),\\ &\gamma_{a1}\tilde{W}_{a1}^T(t)S_1(z_1)S_1^T(z_1)\hat{W}_{a1}(t)\\ &\quad= \frac{\gamma_{a1}}{2}\tilde{W}_{a1}^T(t)S_1(z_1)S_1^T(z_1)\tilde{W}_{a1}(t) + \frac{\gamma_{a1}}{2}\hat{W}_{a1}^T(t)S_1(z_1)\\ &\qquad\times S_1^T(z_1)\hat{W}_{a1}(t) - \frac{\gamma_{a1}}{2}W_1^{*T}S_1(z_1)S_1^T(z_1)W_1^*.\end{aligned}$$

Inserting the above-mentioned results into (28) yields

$$\begin{aligned}&\dot{L}_1(t)\\ &= -\beta_1 z_1^2(t) + z_1(t)z_2(t) - \frac{1}{2}z_1(t)W_1^{*T}S_1(z_1)\\ &\quad-\frac{\gamma_{a1}}{2}\tilde{W}_{a1}^T(t)S_1(z_1)S_1^T(z_1)\tilde{W}_{a1}(t) - \frac{\gamma_{a1}}{2}\hat{W}_{a1}^T(t)\\ &\quad\times S_1(z_1)S_1^T(z_1)\hat{W}_{a1}(t) + \frac{\gamma_{a1}}{2}W_1^{*T}S_1(z_1)S_1^T(z_1)W_1^*\\ &\quad+\frac{\gamma_{c1}}{4(\|\omega_1\|^2+1)}\tilde{W}_{a1}^T(t)S_1(z_1)S_1^T(z_1)\hat{W}_{a1}(t)\omega_1^T\hat{W}_{c1}(t)\\ &\quad-\frac{\gamma_{c1}}{\|\omega_1(t)\|^2+1}\tilde{W}_{c1}^T(t)\omega_1\\ &\quad\times\bigg(\omega_1^T\hat{W}_{c1}(t) - (\beta_1^2-1)z_1^2(t) + 2\beta_1 z_1(t)(f_1(\bar{x}_1) - \dot{y}_r)\\ &\quad+\frac{1}{4}\hat{W}_{a1}^T S_1(z_1)S_1^T(z_1)\hat{W}_{a1}\bigg) + z_1(t)f_1(\bar{x}_1) - z_1(t)\dot{y}_r.\end{aligned} \tag{29}$$

According to Young's inequality $ab \le (a^2/2) + (b^2/2)$, there are the following results:

$$\begin{aligned}z_1(t)z_2(t) &\le z_1^2(t) + z_2^2(t)\\ z_1(t)f_1(\bar{x}_1) &\le \frac{1}{2}z_1^2(t) + \frac{1}{2}f_1^2(\bar{x}_1)\\ -z_1(t)\dot{y}_r(t) &\le \frac{1}{2}z_1^2(t) + \frac{1}{2}\dot{y}_r^2(t)\\ -\frac{1}{2}z_1(t)W_1^{*T}S_1(z_1) &\le z_1^2(t) + \frac{1}{2}\big(W_1^{*T}S_1(z_1)\big)^2.\end{aligned}$$

Adding the above-mentioned inequalities to (29) has

$$\begin{aligned}\dot{L}_1(t) \le& z_2^2(t) - (\beta_1 - 3)z_1^2(t) - \frac{\gamma_{a1}}{2}\tilde{W}_{a1}^T(t)S_1(z_1)\\ &\times S_1^T(z_1)\tilde{W}_{a1}(t) - \frac{\gamma_{a1}}{2}\hat{W}_{a1}^T(t)S_1(z_1)S_1^T(z_1)\hat{W}_{a1}(t)\\ &+\frac{\gamma_{c1}}{4(\|\omega_1\|^2+1)}\tilde{W}_{a1}^T(t)S_1(z_1)S_1^T(z_1)\hat{W}_{a1}(t)\omega_1^T\hat{W}_{c1}(t)\\ &-\frac{\gamma_{c1}}{\|\omega_1\|^2+1}\tilde{W}_{c1}^T(t)\omega_1\\ &\times\bigg(\omega_1^T\hat{W}_{c1}(t) - (\beta_1^2-1)z_1^2(t) + 2\beta_1 z_1(f_1(\bar{x}_1) - \dot{y}_r)\\ &+\frac{1}{4}\hat{W}_{a1}^T(t)S_1(z_1)S_1^T(z_1)\hat{W}_{a1}(t)\bigg)\\ &+\frac{1}{2}f_1^2(\bar{x}_1) + \frac{1}{2}\dot{y}_r^2 + \frac{\gamma_{a1}+1}{2}\big(W_1^{*T}S_1(z_1)\big)^2.\end{aligned} \tag{30}$$

Considering the following equation derived from (17):

$$\begin{aligned}&-(\beta_1^2-1)z_1^2 + 2\beta_1 z_1(f_1(\bar{x}_1) - \dot{y}_r)\\ &\quad= -\omega_1^T W_1^* - \frac{1}{2}\hat{W}_{a1}^T(t)S_1(z_1)S_1^T(z_1)W_1^*\\ &\qquad+\frac{1}{4}W_1^{*T}S_1(z_1)S_1^T(z_1)W_1^* - \epsilon_1(t)\end{aligned} \tag{31}$$

(30) can become

$$\begin{aligned}\dot{L}_1(t) \le& z_2^2(t) - (\beta_1 - 3)z_1^2(t) - \frac{\gamma_{a1}}{2}\tilde{W}_{a1}^T(t)S_1(z_1)\\ &\times S_1^T(z_1)\tilde{W}_{a1}(t) - \frac{\gamma_{a1}}{2}\hat{W}_{a1}^T(t)S_1(z_1)S_1^T(z_1)\hat{W}_{a1}(t)\\ &+\frac{\gamma_{c1}}{4(\|\omega_1\|^2+1)}\tilde{W}_{a1}^T(t)S_1(z_1)S_1^T(z_1)\hat{W}_{a1}(t)\omega_1^T\hat{W}_{c1}(t)\\ &-\frac{\gamma_{c1}}{\|\omega_1(t)\|^2+1}\tilde{W}_{c1}^T(t)\omega_1(t)\\ &\times\bigg(\omega_1^T(t)\tilde{W}_{c1}(t) - \frac{1}{2}\hat{W}_{a1}^T(t)S_1(z_1)S_1^T(z_1)W_1^*\\ &+\frac{1}{4}W_1^{*T}S_1(z_1)S_1^T(z_1)W_1^*\\ &+\frac{1}{4}\hat{W}_{a1}^T(t)S_1(z_1)S_1^T(z_1)\hat{W}_{a1}(t) - \epsilon_1(t)\bigg)\\ &+\frac{1}{2}f_1^2(\bar{x}_1) + \frac{1}{2}\dot{y}_r^2(t) + \frac{\gamma_{a1}+1}{2}\big(W_1^{*T}S_1(z_1)\big)^2.\end{aligned} \tag{32}$$

Inserting the facts that

$$-\frac{1}{2}\hat{W}_{a1}^T(t)S_1(z_1)S_1^T(z_1)W_1^* + \frac{1}{4}W_1^{*T}S_1(z_1)S_1^T(z_1)W_1^* + \frac{1}{4}\hat{W}_{a1}^T(t)S_1(z_1)S_1^T(z_1)\hat{W}_{a1}(t) = \frac{1}{4}\tilde{W}_{a1}^T(t)S_1(z_1)S_1^T(z_1)\hat{W}_{a1}(t) - \frac{1}{4}W_1^{*T}S_1(z_1)S_1^T(z_1)\tilde{W}_{a1}(t),$$

$$\frac{\gamma_{c1}}{\|\omega_1\|^2+1}\tilde{W}_{c1}^T(t)\omega_1(t)\epsilon_1(t) \le \frac{\gamma_{c1}}{2(\|\omega_1\|^2+1)}\tilde{W}_{c1}^T(t)\omega_1(t)\omega_1^T(t)\tilde{W}_{c1}(t) + \frac{\gamma_{c1}}{2(\|\omega_1\|^2+1)}\epsilon_1^2(t) \tag{33}$$

into (32), there is the following one:

$$\begin{aligned}\dot{L}_1(t) \le& -(\beta_1-3)z_1^2(t) + z_2^2(t) - \frac{\gamma_{a1}}{2}\tilde{W}_{a1}^T(t)S_1(z_1) \times S_1^T(z_1)\tilde{W}_{a1}(t) - \frac{\gamma_{c1}}{2(\|\omega_1\|^2+1)}\tilde{W}_{c1}^T(t)\omega_1\omega_1^T\tilde{W}_{c1}(t)\\ &+\frac{\gamma_{c1}}{4(\|\omega_1\|^2+1)}\tilde{W}_{a1}^T(t)S_1(z_1)S_1^T(z_1)\hat{W}_{a1}(t)\omega_1^T\hat{W}_{c1}(t)\\ &-\frac{\gamma_{c1}}{4(\|\omega_1\|^2+1)}\tilde{W}_{c1}^T(t)\omega_1\tilde{W}_{a1}^T(t)S_1(z_1)S_1^T(z_1)\hat{W}_{a1}(t)\\ &+\frac{\gamma_{c1}}{4(\|\omega_1\|^2+1)}\tilde{W}_{c1}^T(t)\omega_1W_1^{*T}S_1(z_1)S_1^T(z_1)\tilde{W}_{a1}(t)\\ &-\frac{\gamma_{a1}}{2}\hat{W}_{a1}^T(t)S_1(z_1)S_1^T(z_1)\hat{W}_{a1}(t) + \frac{1}{2}f_1^2(\bar{x}_1) + \frac{1}{2}\dot{y}_r^2\\ &+\frac{\gamma_{a1}+1}{2}\left(W_1^{*T}S_1(z_1)\right)^2 + \frac{\gamma_{c1}}{2}\epsilon_1^2(t).\end{aligned} \tag{34}$$

Substituting the following fact:

$$\begin{aligned}&\frac{\gamma_{c1}}{4(\|\omega_1\|^2+1)}\tilde{W}_{a1}^T(t)S_1(z_1)S_1^T(z_1)\hat{W}_{a1}(t)\omega_1^T\hat{W}_{c1}(t) - \frac{\gamma_{c1}}{4(\|\omega_1\|^2+1)}\tilde{W}_{c1}^T(t)\omega_1\tilde{W}_{a1}^T(t)S_1(z_1)S_1^T(z_1)\hat{W}_{a1}(t)\\ &= \frac{\gamma_{c1}}{4(\|\omega_1\|^2+1)}\tilde{W}_{a1}^T(t)S_1(z_1)\hat{W}_{c1}^T(t)\omega_1S_1^T(z_1)\hat{W}_{a1}(t) - \frac{\gamma_{c1}}{4(\|\omega_1\|^2+1)}\tilde{W}_{a1}^T(t)S_1(z_1)\tilde{W}_{c1}^T(t)\omega_1S_1^T(z_1)\hat{W}_{a1}(t)\\ &= \frac{\gamma_{c1}}{4(\|\omega_1\|^2+1)}\tilde{W}_{a1}^T(t)S_1(z_1)W_1^{*T}\omega_1S_1^T(z_1)\hat{W}_{a1}(t)\end{aligned}$$

into inequality (34) yields

$$\begin{aligned}\dot{L}_1(t) \le& -(\beta_1-3)z_1^2(t) + z_2^2(t) - \frac{\gamma_{a1}}{2}\tilde{W}_{a1}^T(t)S_1(z_1) \times S_1^T(z_1)\tilde{W}_{a1}(t) - \frac{\gamma_{c1}}{2(\|\omega_1\|^2+1)}\tilde{W}_{c1}^T(t)\omega_1\omega_1^T\tilde{W}_{c1}(t)\\ &+\frac{\gamma_{c1}}{4(\|\omega_1\|^2+1)}\tilde{W}_{a1}^T(t)S_1(z_1)W_1^{*T}\omega_1S_1^T(z_1)\hat{W}_{a1}(t)\\ &+\frac{\gamma_{c1}}{4(\|\omega_1\|^2+1)}\tilde{W}_{c1}^T(t)\omega_1W_1^{*T}S_1(z_1)S_1^T(z_1)\tilde{W}_{a1}(t)\\ &-\frac{\gamma_{a1}}{2}\hat{W}_{a1}^T(t)S_1(z_1)S_1^T(z_1)\hat{W}_{a1}(t) + \frac{1}{2}f_1^2(\bar{x}_1) + \frac{1}{2}\dot{y}_r^2\\ &+\frac{\gamma_{a1}+1}{2}\left(W_1^{*T}S_1(z_1)\right)^2 + \frac{\gamma_{c1}}{2}\epsilon_1^2(t).\end{aligned} \tag{35}$$

According to Young's inequality and Cauchy inequality, there are the following results:

$$\begin{aligned}&\frac{\gamma_{c1}}{4(\|\omega_1\|^2+1)}\tilde{W}_{a1}^T(t)S_1(z_1)W_1^{*T}\omega_1(t)S_1^T(z_1)\hat{W}_{a1}(t)\\ &\le \frac{1}{32}\tilde{W}_{a1}^T(t)S_1(z_1)W_1^{*T}\omega_1\omega_1^TW_1^*S_1^T(z_1)\tilde{W}_{a1}(t) + \frac{\gamma_{c1}^2}{2}\hat{W}_{a1}^T(t)S_1(z_1)S_1^T(z_1)\hat{W}_{a1}(t),\\ &\frac{\gamma_{c1}}{4(\|\omega_1\|^2+1)}\tilde{W}_{c1}^T(t)\omega_1(t)W_1^{*T}S_1(z_1)S_1^T(z_1)\tilde{W}_{a1}(t)\\ &\le \frac{\gamma_{c1}^2}{2}\tilde{W}_{a1}^T(t)S_1(z_1)S_1^T(z_1)\tilde{W}_{a1}(t) + \frac{1}{32(\|\omega_1\|^2+1)} \times \tilde{W}_{c1}^T(t)\omega_1W_1^{*T}S_1(z_1)S_1^T(z_1)W_1^*\omega_1^T\tilde{W}_{c1}(t).\end{aligned}$$

Applying the above-mentioned inequalities to (35) has

$$\begin{aligned}\dot{L}_1(t) \le& z_2^2(t) - (\beta_1-3)z_1^2(t) - \left(\frac{\gamma_{a1}}{2} - \frac{\gamma_{c1}^2}{2} - \frac{1}{32}W_1^{*T}\omega_1\omega_1^TW_1^*\right) \times \tilde{W}_{a1}^T(t)S_1(z_1)S_1^T(z_1)\tilde{W}_{a1}(t)\\ &-\frac{1}{\|\omega_1\|^2+1}\left(\frac{\gamma_{c1}}{2} - \frac{1}{32}W_1^{*T}S_1(z_1)S_1^T(z_1)W_1^*\right) \times\tilde{W}_{c1}^T(t)\omega_1\omega_1^T\tilde{W}_{c1}(t) - \left(\frac{\gamma_{a1}}{2} - \frac{\gamma_{c1}^2}{2}\right)\hat{W}_{a1}^T(t)\\ &\times S_1(z_1)S_1^T(z_1)\hat{W}_{a1}(t) + \frac{1}{2}f_1^2(\bar{x}_1) + \frac{1}{2}\dot{y}_r^2(t) + \frac{\gamma_{a1}+1}{2}\left(W_1^{*T}S_1(z_1)\right)^2 + \frac{\gamma_{c1}}{2}\epsilon_1^2(t).\end{aligned} \tag{36}$$

Rewrite (36) to compact form as

$$\dot{L}_1(t) \le -\xi_1^T(t)A_1(t)\xi_1(t) + C_1(t) + z_2^2(t) - \left(\frac{\gamma_{a1}}{2} - \frac{\gamma_{c1}^2}{2}\right)\hat{W}_{a1}^T(t)S_1(z_1)S_1^T(z_1)\hat{W}_{a1}(t) \tag{37}$$

where $\xi_1(t)$, $A_1(t)$, and C_1 are shown at the top of the next page.

Because the PE condition is held (Assumption 1), the diagonal matrix $A_1(t)$ can be made positive definite by designing the parameters β_1, γ_{c1}, and γ_{a1} to satisfy the following conditions:

$$\beta_1 > 3, \quad \gamma_{a1} > \gamma_{c1}^2 + \frac{\zeta_1}{16}W_1^{*T}W_1^*$$
$$\gamma_{c1} > \frac{1}{16}\sup_{t\ge0}\left\{\lambda_{\max}\left\{W_1^{*T}S_1(z_1)S_1^T(z_1)W_1^*\right\}\right\}. \tag{38}$$

Remark 2 It should be mentioned that the unknown constant matrix W_1^* is only for analysis purpose. The condition (38) implies that the matrix $A_1(t)$ can be made positive definition.

Then, (37) can become the following one:

$$\dot{L}_1(t) < -a_1\|\xi_1(t)\|^2 + c_1 + z_2^2(t)$$

where $a_1 = \inf_{t\ge0}\{\lambda_{\min}\{A_1(t)\}\}$, $c_1 = \sup_{t\ge0}\{C_1(t)\}$.

Step i ($i = 2, \dots, n-1$): Define the tracking error variable for the ith step as $z_i(t) = x_i(t) - \hat{\alpha}_{i-1}(z_{i-1})$. Based on the system dynamic (8), the error dynamic for z_i-subsystem is

$$\dot{z}_i(t) = f_i(\bar{x}_i) + x_{i+1}(t) - \dot{\hat{\alpha}}_{i-1}(z_{i-1}). \tag{39}$$

$$\xi_1(t) = \left[z_1(t), \tilde{W}_{a1}^T(t), \tilde{W}_{c1}^T(t)\right]^T$$

$$A_1(t) = \begin{bmatrix} \beta_1 - 3 & 0 & 0 \\ 0 & \left(\frac{\gamma_{a1}}{2} - \frac{\gamma_{c1}^2}{2} - \frac{1}{32} W_1^{*T} \omega_1 \omega_1^T W_1^*\right) S_1(z_1) S_1^T(z_1) & 0 \\ 0 & 0 & \frac{1}{\|\omega_1(t)\|^2 + 1}\left(\frac{\gamma_{c1}}{2} - \frac{1}{32} W_1^{*T} S_1(z_1) S_1^T(z_1) W_1^*\right)\omega_1(t)\omega_1^T(t) \end{bmatrix}$$

$$C_1 = \frac{1}{2} f_1^2(\bar{x}_1) + \frac{1}{2}\dot{y}_r^2(t) + \frac{\gamma_{a1}+1}{2}\left(W_1^{*T} S_1(z_1)\right)^2 + \frac{\gamma_{c1}}{2}\epsilon_1^2(t)$$

Viewing $x_{i+1}(t)$ as the optimal virtual control input $\alpha_i^*(z_i)$, the optimal value function for z_i-subsystem is defined as

$$V_i^*(z_i) = \min_{\alpha_i \in \Psi(\Omega_{z_i})} \left(\int_t^\infty r_i(z_i(s), \alpha_i(z_i)) ds\right) = \int_t^\infty r_i\left(z_i(s), \alpha_i^*(z_i)\right) ds$$

where $r_i(z_i, \alpha_i) = z_i^2(t) + \alpha_i^2(z_i)$ is the cost function, α_i is the virtual controller, and Ω_{z_i} is a compact set containing origin. The optimal value function $V_i^*(z_i)$ is reexpressed as

$$V_i^*(z_i) = \beta_i z_i^2(t) + V_i^o(z_i)$$

where β_i is a positive constant and $V_i^o(z_i) = -\beta_i z_i^2(t) + V_i^*(z_i)$ is a continuous scalar function.

The HJB equation for z_i-subsystem is

$$H_i\left(z_i, \alpha_i^*, \frac{\partial V_i^*}{\partial z_i}\right) = z_i^2 + \alpha_i^{*2} + \left(2\beta_i z_i + \frac{\partial V_i^o(z_i)}{\partial z_i}\right) \times \left(f_i(\bar{x}_i) + \alpha_i^*(z_i) - \dot{\hat{\alpha}}_{i-1}(z_{i-1})\right) = 0. \tag{40}$$

Then, the optimal controller $\alpha_i^*(z_i)$ can be get by solving the equation $\partial H_i / \partial \alpha_i^* = 0$

$$\alpha_i^*(z_i) = -\beta_i z_i(t) - \frac{1}{2}\frac{\partial V_i^o(z_i)}{\partial z_i}. \tag{41}$$

For any $z_i \in \Omega_{z_i}$, $\partial V_i^o(z_i)/\partial z_i$ can be approximated by NN as

$$\frac{\partial V_i^o(z_i)}{\partial z_i} = W_i^{*T} S_i(z_i) + \varepsilon_i(z_i) \tag{42}$$

where $W_i^{*T} \in R^{m_i}$ is the ideal weight, $S_i(z_i) \in R^{m_i}$ is the basis function vector, and $\varepsilon_i(z_i) \in R$ is the approximation error, which is bounded by a constant δ_i, i.e., $|\varepsilon_i(z_i)| \le \delta_i$.

The gradient term $\partial V_i^*(z_i)/\partial z_i$ and the optimal controller $\alpha_i^*(z_i)$ can be redescribed as

$$\frac{\partial V_i^*(z_i)}{\partial z_i} = 2\beta_i z_i(t) + W_i^{*T} S_i(z_i) + \varepsilon_i(z_i) \tag{43}$$

$$\alpha_i^*(z_i) = -\beta_i z_i(t) - \frac{1}{2}\left(W_i^{*T} S_i(z_i) + \varepsilon_i(z_i)\right). \tag{44}$$

Substituting (42) and (44) into (40), the following one can be obtained:

$$H_i\left(z_i, \alpha_i^*, W_i^*\right) = -(\beta_i^2 - 1) z_i^2(t) + 2\beta_i z_i(t)(f_i(\bar{x}_i) - \dot{\hat{\alpha}}_{i-1}) + W_i^{*T} S_i(z_i)(f_i(\bar{x}_i) - \dot{\hat{\alpha}}_{i-1} - \beta_i z_i(t)) - \frac{1}{4} W_i^{*T} S_i(z_i) S_i^T(z_i) W_i^* + \epsilon_i(t) = 0 \tag{45}$$

where $\epsilon_i(t) = \varepsilon_i(z_i)(f_i(\bar{x}_i) - \dot{\hat{\alpha}}_{i-1} + \alpha_i^*) + (1/4)\varepsilon_i^2(z_i)$, which is bounded because all terms are bounded.

The optimal controller (41) is unavailable, because the ideal weight W_i^* is unknown, and in order to obtain the valid controller, the actor–critic RL is used, where the critic and actor NNs are designed as

$$\frac{\partial \hat{V}_i^*(z_i)}{\partial z_i} = 2\beta_i z_i(t) + \frac{\partial \hat{V}_i^o(z_i)}{\partial z_i} = 2\beta_i z_i(t) + \hat{W}_{ci}^T(t) S_i(z_i)$$

$$\hat{\alpha}_i(z_i) = -\beta_i z_i(t) - \frac{1}{2}\hat{W}_{ai}^T(t) S_i(z_i) \tag{46}$$

where $\hat{V}_i^*(z_i)$ and $\hat{V}_i^o(z_i)$ are the estimations of $V_i^*(z_i)$ and $V_i^o(z_i)$, respectively; $\hat{W}_{ci}^T(t) \in R^{m_i}$ and $\hat{W}_{ai}^T(t) \in R^{m_i}$ are the critic and actor NN weights, respectively.

Remark 3: The boundedness of $\dot{\hat{\alpha}}_{i-1}$, $i = 1, \ldots, n$, is proven in the following.

Based on (39) and (46), the time derivative $\dot{\hat{\alpha}}_{i-1}$, $i = 2, \ldots, n$, can be expressed as

$$\dot{\hat{\alpha}}_{i-1}(t) = -\beta_{i-1}(f_{i-1}(\bar{x}_{i-1}) + x_i(t) - \dot{\hat{\alpha}}_{i-2}(z_{i-2})) - \frac{1}{2}\left(\dot{\hat{W}}_{a(i-1)}^T S_{i-1}(z_{i-1}) + \hat{W}_{a(i-1)}^T \dot{S}_{i-1}(z_{i-1})\right). \tag{47}$$

Since these terms $f_{i-1}(\bar{x}_{i-1}) + x_i(t)$, $i = 1, \ldots, n$, are Lipschitz continuous, they are bounded for $z_i \in \Omega_{z_i}$. Starting from $\dot{\hat{\alpha}}_1 = -\beta_1(f_1(\bar{x}_1) + x_2(t) - \dot{y}_r) - (1/2)(\dot{\hat{W}}_{a1}^T(t) S_1(z_1) + \hat{W}_{a1}^T(t)\dot{S}_1(z_1))$ that is bounded, it can be successively proved that $\dot{\hat{\alpha}}_{i-1}$, $i = 3, \ldots, n$, are bounded by using (47). □

The approximated HJB equation for z_i-subsystem is

$$H_i(z_i, \hat{\alpha}_i, \hat{W}_{ci}) = z_i^2(t) + \left(-\beta_i z_i(t) - \frac{1}{2}\hat{W}_{ai}^T(t) S_i(z_i)\right)^2 + \left(2\beta_i z_i(t) + \hat{W}_{ci}^T(t) S_i(z_i)\right)\left(-\beta_i z_i(t) - \frac{1}{2}\hat{W}_{ai}^T(t) S_i(z_i) + f_i(\bar{x}_i) - \dot{\hat{\alpha}}_{i-1}(z_{i-1})\right).$$

The Bellman residual error is yielded as $e_i(t) = H_i(z_i, \hat{\alpha}_i, \hat{W}_{ci})$. Define the positive definite function as $E_i(t) = (1/2)e_i^2(t)$. The critic NN weight updating law is designed by using gradient descent algorithm

$$\begin{aligned}\dot{\hat{W}}_{ci}(t) &= -\frac{\gamma_{ci}}{\|\omega_i(t)\|^2+1}e_i(t)\frac{\partial e_i(t)}{\partial \hat{W}_{ci}(t)}\\ &= -\frac{\gamma_{ci}}{\|\omega_i(t)\|^2+1}\omega_i(t)\\ &\quad\times\Big(\omega_i^T(t)\hat{W}_{ci}(t)-(\beta_i^2-1)z_i^2(t)\\ &\quad+2\beta_i z_i(f_i(\bar{x}_i)-\dot{\hat{\alpha}}_{i-1})+\frac{1}{4}\hat{W}_{ai}^T S_i(z_i)S_i^T(z_i)\hat{W}_{ai}\Big)\end{aligned} \tag{48}$$

where $\gamma_{ci} > 0$ is the learning rate and $\omega_i(t) = S_i(z_i)(f_i(\bar{x}_i) - \dot{\hat{\alpha}}_{i-1} - \beta_i z_i(t) - (1/2)\hat{W}_{ai}^T(t)S_i(z_i)) \in R^{m_i}$.

The actor NN weight law is designed based on the stability analysis

$$\begin{aligned}\dot{\hat{W}}_{ai}(t) &= \frac{1}{2}S_i(z_i)z_i(t) - \gamma_{ai}S_i(z_i)S_i^T(z_i)\hat{W}_{ai}(t)\\ &\quad+\frac{\gamma_{ci}}{4(\|\omega_i\|^2+1)}S_i(z_i)S_i^T(z_i)\hat{W}_{ai}(t)\omega_i^T(t)\hat{W}_{ci}(t)\end{aligned} \tag{49}$$

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

Using the error variable $z_{i+1}(t) = x_{i+1}(t) - \hat{\alpha}_i(z_i)$, the error dynamic (39) can be rewritten as

$$\dot{z}_i(t) = f_i(\bar{x}_i(t)) + z_{i+1}(t) + \hat{\alpha}_i(z_i) - \dot{\hat{\alpha}}_{i-1}(z_{i-1}). \tag{50}$$

Consider the following Lyapunov function candidate for z_i-subsystem:

$$L_i(t) = \sum_{k=1}^{i-1}L_k(t) + \frac{1}{2}z_i^2 + \frac{1}{2}\tilde{W}_{ai}^T(t)\tilde{W}_{ai}(t) + \frac{1}{2}\tilde{W}_{ci}^T(t)\tilde{W}_{ci}(t)$$

where $\tilde{W}_{ci}(t) = \hat{W}_{ci}(t) - W_i^*$ and $\tilde{W}_{ai}(t) = \hat{W}_{ai}(t) - W_i^*$ are the critic and actor NN estimation errors.

Based on (48), (49), and (50), the time derivative of $L_i(t)$ is

$$\begin{aligned}&\dot{L}_i(t)\\ &= \sum_{k=1}^{i-1}\dot{L}_k(t) + z_i(f_i(\bar{x}_i) + z_{i+1} - \dot{\hat{\alpha}}_{i-1} + \hat{\alpha}_i)\\ &\quad-\tilde{W}_{ai}^T(t)\Big(\frac{1}{2}S_i(z_i)z_i(t) + \gamma_{ai}S_i(z_i)S_i^T(z_i)\hat{W}_{ai}(t)\\ &\quad-\frac{\gamma_{ci}}{4(\|\omega_i\|^2+1)}S_i(z_i)S_i^T(z_i)\hat{W}_{ai}(t)\omega_i^T(t)\hat{W}_{ci}(t)\Big)\\ &\quad-\frac{\gamma_{ci}}{\|\omega_i\|^2+1}\tilde{W}_{ci}^T(t)\omega_i\\ &\quad\times\Big(\omega_i^T\hat{W}_{ci}(t) - (\beta_i^2-1)z_i^2(t)\\ &\quad+2\beta_i z_i(t)(f_i(\bar{x}_i) - \dot{\hat{\alpha}}_{i-1}(z_{i-1})) + \frac{1}{4}\hat{W}_{ai}^T(t)S_i(z_i)\\ &\quad\times S_i^T(z_i)\hat{W}_{ai}(t)\Big).\end{aligned}$$

Similar to step 1, the following result can be derived:

$$\begin{aligned}\dot{L}_i(t) &\le \sum_{k=1}^{i-1}\dot{L}_k(t) + z_{i+1}^2 - (\beta_i - 3)z_i^2\\ &\quad-\left(\frac{\gamma_{ai}}{2} - \frac{\gamma_{ci}^2}{2} - \frac{1}{32}W_i^{*T}\omega_i\omega_i^T W_i^*\right)\\ &\quad\times\tilde{W}_{ai}^T(t)S_i(z_i)S_i^T(z_i)\tilde{W}_{ai}(t)\\ &\quad-\frac{1}{\|\omega_i\|^2+1}\left(\frac{\gamma_{ci}}{2} - \frac{1}{32}W_i^{*T}S_i(z_i)S_i^T(z_i)W_i^*\right)\\ &\quad\times\tilde{W}_{ci}^T(t)\omega_i(t)\omega_i^T(t)\tilde{W}_{ci}(t) - \left(\frac{\gamma_{ai}}{2} - \frac{\gamma_{ci}^2}{2}\right)\hat{W}_{ai}^T(t)\\ &\quad\times S_i(z_i)S_i^T(z_i)\hat{W}_{ai}(t) + \frac{1}{2}f_i^2(\bar{x}_i) + \frac{1}{2}\dot{\hat{\alpha}}_{i-1}^2(z_{i-1})\\ &\quad+\frac{\gamma_{ai}+1}{2}\big(W_i^{*T}S_i(z_i)\big)^2 + \frac{\gamma_{ci}}{2}\epsilon_i^2(t).\end{aligned} \tag{51}$$

Using results of the first $i-1$ steps, the inequality (51) is rewritten as

$$\begin{aligned}\dot{L}_i(t) &\le \sum_{k=1}^{i-1}\Big(-a_k\|\xi_k(t)\|^2 + c_k\Big) + z_{i+1}^2(t) - \xi_i^T(t)A_i(t)\xi_i(t)\\ &\quad+C_i(t) - \left(\frac{\gamma_{ai}}{2} - \frac{\gamma_{ci}^2}{2}\right)\hat{W}_{ai}^T(t)S_i(z_i)S_i^T(z_i)\hat{W}_{ai}(t)\end{aligned}$$

where $\xi_i(t)$, $A_i(t)$, and $C_i(t)$ are shown at the top of the next page.

Based on Assumption 1, the matrix $A_i(t)$ can be positive definite by choosing the parameters β_i, γ_{ci}, and γ_{ai} satisfying the following conditions:

$$\begin{aligned}&\beta_i > 4,\quad \gamma_{ai} > \gamma_{ci}^2 + \frac{\zeta_i}{16}W_i^{*T}W_i^*\\ &\gamma_{ci} > \frac{1}{16}\sup_{t\ge 0}\big\{\lambda_{\max}\big\{W_i^{*T}S_i(z_i)S_i^T(z_i)W_i^*\big\}\big\}.\end{aligned} \tag{52}$$

Then, there is the following inequality:

$$\dot{L}_i(t) < \sum_{k=1}^{i}(-a_k\|\xi_k(t)\|^2 + c_k) + z_{i+1}^2(t)$$

where $a_k = \inf_{t\ge 0}\{\lambda_{\min}\{A_k(t)\}\}$ and $c_k = \sup_{t\ge 0}\{C_k(t)\}$.

Step n: In the final step of the backstepping control, the actual controller u will be derived. Defining the tracking error variable for the nth step as $z_n(t) = x_n(t) - \hat{\alpha}_{n-1}(z_{n-1})$, based on the system dynamic (8), the error dynamic is

$$\dot{z}_n(t) = f_n(\bar{x}_n(t)) + u - \dot{\hat{\alpha}}_{n-1}(z_{n-1}). \tag{53}$$

The optimal value function is defined as

$$\begin{aligned}V_n^*(z_n) &= \min_{u\in\Psi(\Omega_{z_n})}\left(\int_t^\infty r_n(z_i(s), u(z_n))ds\right)\\ &= \int_t^\infty r_n(z_n(s), u^*(z_n))ds\end{aligned}$$

where $r_n(z_n, u) = z_n^2(t) + u^2$ is the cost function, u^* is the optimal actual control, and Ω_{z_n} is a compact set containing origin. Rewrite the optimal value function as

$$V_n^*(z_n) = \beta_n z_n^2(t) + V_n^o(z_n) \tag{54}$$

where β_n is a positive constant and $V_n^o(z_n) = -\beta_n z_n^2(t) + V_n^*(z_n)$ is a continuous scalar function.

The HJB equation for the subsystem is

$$\begin{aligned}H_n\left(z_n, u^*, \frac{\partial V_n^*}{\partial z_n}\right) &= z_n^2(t) + u^{*2} + \left(2\beta_n z_n(t) + \frac{\partial V_n^o}{\partial z_n}\right)\\ &\quad\times(f_n(\bar{x}_n) - \dot{\hat{\alpha}}_{n-1}(z_{n-1}) + u^*) = 0.\end{aligned} \tag{55}$$

$$\xi_i(t) = \left[z_i(t), \tilde{W}_{ai}^T(t), \tilde{W}_{ci}^T(t)\right]^T$$

$$A_i(t) = \begin{bmatrix} \beta_i - 4 & 0 & 0 \\ 0 & \left(\frac{\gamma_{ai}}{2} - \frac{\gamma_{ci}^2}{2} - \frac{1}{32} W_i^{*T}\omega_i(t)\omega_i^T(t)W_i^*\right) S_i(z_i)S_i^T(z_i) & 0 \\ 0 & 0 & \frac{1}{\|\omega_i(t)\|^2+1}\left(\frac{\gamma_{ci}}{2} - \frac{1}{32}W_i^{*T}S_i(z_i)S_i^T(z_i)W_i^*\right)\omega_i(t)\omega_i^T(t) \end{bmatrix}$$

$$C_i(t) = \frac{\gamma_{ai}+1}{2}\left(W_i^{*T}S_i(z_i)\right)^2 + \frac{1}{2}f_i^2(\bar{x}_i) + \frac{1}{2}\dot{\hat{\alpha}}_{i-1}^2 + \frac{\gamma_{ci}}{2}\epsilon_i^2(t)$$

The optimal control $u^*(z_n)$ can be got by solving $\partial H_n/\partial u^* = 0$

$$u^*(z_n) = -\beta_n z_n(t) - \frac{1}{2}\frac{\partial V_n^o(z_n)}{\partial z_n}. \quad (56)$$

For any $z_n \in \Omega_{z_n}$, the uncertain term $\partial V_n^o(z_n(t))/\partial z_n$ is approximated as

$$\frac{\partial V_n^o(z_n)}{\partial z_n} = W_n^{*T}S_n(z_n) + \varepsilon_n(z_n) \quad (57)$$

where $W_n^* \in R^{m_n}$ is the ideal weight, $S_n(z_n) \in R^{m_n}$ is the basis function vector, $\varepsilon_n(z_n) \in R$ is the approximation error satisfying $|\varepsilon_n(z_n)| \le \delta_n$, and δ_n is a positive constant.

The gradient term $\partial V_n^*(z_n)/\partial z_n$ and the optimal controller $u^*(z_n)$ can be redescribed as

$$\frac{\partial V_n^*(z_n)}{\partial z_n} = 2\beta_n z_n(t) + W_n^{*T}S_n(z_n) + \varepsilon_n(z_n) \quad (58)$$

$$u^*(z_n) = -\beta_n z_n(t) - \frac{1}{2}\left(W_n^{*T}S_n(z_n) + \varepsilon_n\right). \quad (59)$$

Substituting (57) and (59) into (55), the following one can be obtained:

$$\begin{aligned} H_n(z_n, u^*, W_n^*) &= -\left(\beta_n^2 - 1\right)z_n^2(t) + 2\beta_n z_n(t)(f_n(\bar{x}_n) - \dot{\hat{\alpha}}_{n-1}) \\ &\quad + W_n^{*T}S_n(z_n)(f_n(\bar{x}_n) - \dot{\hat{\alpha}}_{n-1} - \beta_n z_n(t)) \\ &\quad - \frac{1}{4}W_n^{*T}S_n(z_n)S_n^T(z_n)W_n^* + \epsilon_n(t) = 0 \end{aligned}$$

where $\epsilon_n(t) = \varepsilon_n(z_n)(f_n(\bar{x}_n) - \dot{\hat{\alpha}}_{n-1} + u^*) + (1/4)\varepsilon_n^2(z_n)$, which is a bounded term.

The following critic and actor NNs are designed to implement the RL iteration for the optimized control:

$$\frac{\partial \hat{V}_n^*(z_n)}{\partial z_n} = 2\beta_n z_n + \frac{\partial \hat{V}_n^o(z_n)}{\partial z_n} = 2\beta_n z_n(t) + \hat{W}_{cn}^T(t)S_n(z_n)$$

$$u = -\beta_n z_n(t) - \frac{1}{2}\hat{W}_{an}^T(t)S_i(z_n) \quad (60)$$

where $\hat{V}_n^*(z_n)$ and $\hat{V}_n^o(z_n)$ are the estimations of $V_n^*(z_n)$ and $V_n^o(z_n)$, respectively; $\hat{W}_{cn}^T(t) \in R^{m_n}$ and $\hat{W}_{an}^T(t) \in R^{m_n}$ are the critic and actor NN weights, respectively.

The approximated HJB equation is

$$\begin{aligned} &H_n(z_n, u, \hat{W}_{cn}) \\ &= z_n^2 + \left(-\beta_n z_n - \frac{1}{2}\hat{W}_{an}^T(t)S_n(z_n)\right)^2 \\ &\quad + \left(2\beta_n z_n + \hat{W}_{cn}^T(t)S_n(z_n)\right) \\ &\quad \times \left(f_n(\bar{x}_n) - \dot{\hat{\alpha}}_{n-1} - \beta_n z_n(t) - \frac{1}{2}\hat{W}_{an}^T(t)S_n(z_n)\right). \end{aligned}$$

The Bellman residual error is $e_n(t) = H_n(z_n, u, \hat{W}_{cn})$. Defining the positive definite function as $E_n(t) = (1/2)e_n^2(t)$, the following critic NN weight updating law is derived by using the gradient descent algorithm:

$$\begin{aligned} \dot{\hat{W}}_{cn}(t) &= -\frac{\gamma_{cn}}{\|\omega_n(t)\|^2+1}e_n(t)\frac{\partial e_n(t)}{\partial \hat{W}_{cn}(t)} \\ &= -\frac{\gamma_{cn}}{\|\omega_n(t)\|^2+1}\omega_n \\ &\quad \times \Bigg(\omega_n^T\hat{W}_{cn}(t) - (\beta_n - 1)z_n^2(t) \\ &\qquad + 2\beta_n z_n(t)(f_n(\bar{x}_n) - \dot{\hat{\alpha}}_{n-1}) + \frac{1}{4}\hat{W}_{an}^T(t)S_n(z_n) \\ &\qquad \times S_n^T(z_n)\hat{W}_{an}(t)\Bigg) \end{aligned} \quad (61)$$

where $\gamma_{cn} > 0$ are the learning rate and $\omega_n(t) = S_n(z_n)(f_n(\bar{x}_n) - \dot{\hat{\alpha}}_{n-1}(z_{n-1}) - \beta_n z_n(t) - (1/2)\hat{W}_{an}^T(t)S_n(z_n))$.

The actor weight updating law based on the stability analysis is given in the following:

$$\begin{aligned} \dot{\hat{W}}_{an}(t) &= \frac{1}{2}S_n(z_n)z_n(t) - \gamma_{an}S_n(z_n)S_n^T(z_n)\hat{W}_{an}(t) \\ &\quad + \frac{\gamma_{cn}}{4(\|\omega_n(t)\|^2+1)}S_n(z_n)S_n^T(z_n)\hat{W}_{an}\omega_n^T\hat{W}_{cn} \end{aligned} \quad (62)$$

where $\gamma_{an} > 0$ are the learning rate.

Consider the overall Lyapunov function candidate for the final step as

$$L(t) = \sum_{k=1}^{n-1}L_k(t) + \frac{1}{2}z_n^2 + \frac{1}{2}\tilde{W}_{an}^T(t)\tilde{W}_{an}(t) + \frac{1}{2}\tilde{W}_{cn}^T(t)\tilde{W}_{cn}(t)$$

where $\tilde{W}_{cn}(t) = \hat{W}_{cn}(t) - W_n^*$ and $\tilde{W}_{an}(t) = \hat{W}_{an}(t) - W_n^*$ are the critic and actor NN estimation errors, respectively.

Similar to the first $n-1$ steps, the time derivative of $L(t)$ along (53), (61) and (62) satisfies

$$\begin{aligned} \dot{L}(t) &\le \sum_{k=1}^{n-1}\dot{L}_k(t) - (\beta_n - 3)z_n^2(t) \\ &\quad - \left(\frac{\gamma_{an}}{2} - \frac{\gamma_{cn}^2}{2} - \frac{1}{32}W_n^{*T}\omega_n\omega_n^TW_n^*\right) \\ &\quad \times \tilde{W}_{an}^T(t)S_n(z_n)S_n^T(z_n)\tilde{W}_{an}(t) \\ &\quad - \frac{1}{\|\omega_n\|^2+1}\left(\frac{\gamma_{ci}}{2} - \frac{1}{32}W_n^{*T}S_n(z_n)S_n^T(z_n)W_n^*\right) \\ &\quad \times \tilde{W}_{cn}^T(t)\omega_n\omega_n^T\tilde{W}_{cn}(t) - \left(\frac{\gamma_{an}}{2} - \frac{\gamma_{cn}^2}{2}\right)\hat{W}_{an}^T(t) \end{aligned}$$

$$\times S_n(z_n)S_n^T(z_n)\hat{W}_{an}(t) + \frac{1}{2}f_n^2(\bar{x}_n) + \frac{1}{2}\dot{\hat{\alpha}}_{n-1}^2 + \frac{\gamma_{an}+1}{2}\left(W_n^{*T}S_n(z_n)\right)^2 + \frac{\gamma_{cn}}{2}\epsilon_n^2(t). \quad (63)$$

By using the results of previous steps, the inequality (63) is rewritten as

$$\dot{L}(t) \le \sum_{k=1}^{n-1}\left(-a_k\|\xi_k(t)\|^2 + c_k\right) - \xi_n^T(t)A_n(t)\xi_n(t) + C_n(t) - \left(\frac{\gamma_{an}}{2} - \frac{\gamma_{cn}^2}{2}\right)\hat{W}_{an}^T(t)S_n(z_n)S_n^T(z_n)\hat{W}_{an}(t)$$

where $\xi_n(t)$, $A_n(t)$, and $C_n(t)$ are shown at the bottom of the this page.

Based on the PE assumption, the matrix $A_n(t)$ can be made positive definite by satisfying the following conditions:

$$\beta_n > 4, \quad \gamma_{an} > \gamma_{cn}^2 + \frac{\zeta_n}{16}W_n^{*T}W_n^*$$
$$\gamma_{cn} > \frac{1}{16}\sup_{t\ge 0}\left\{\lambda_{\max}\left\{W_n^{*T}S_n(z_n)S_n^T(z_n)W_n^*\right\}\right\}. \quad (64)$$

Let $a_n = \inf_{t\ge 0}\{\lambda_{\min}\{A_n(t)\}\}$ and $c_n = \sup_{t\ge 0}\{C_n(t)\}$, the following result can be obtained:

$$\dot{L}(t) < \sum_{k=1}^{n}(-a_k\|\xi_k(t)\|^2 + c_k). \quad (65)$$

The main results are summarized by the following theorem.

Theorem 1 Consider the strict-feedback system (8) with bounded initial states and reference signal. The control laws choose (60) as the actual control and (19) and (46) as the virtual controls; the weight updating laws are provided by (23), (48), and (61) for critic NNs and (24), (49), and (62) for actor NNs with bounded initial values. If Assumption 1 is held and the design parameters satisfy the conditions (38), (52), and (64), then the optimized high system control scheme can guarantee the following.

1) The error signals $z_i(t)$, $\tilde{W}_{ci}(t)$, and $\tilde{W}_{ai}(t)$ are SGUUB.
2) The desired tracking performance can be obtained.

Proof: 1) The inequality (65) can become as

$$\dot{L}(t) < -aL(t) + c$$

where $a = \min\{a_1, a_2, \ldots, a_n\}$ and $c = \sum_{k=1}^{n} c_k$.

According to Lemma 1, there is the following fact that:

$$L(t) < e^{-at}L(0) + \frac{c}{a}(1 - e^{-at}).$$

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

2) Let $L_z(t) = (1/2)\sum_{k=1}^{n} z_k^2(t)$, the time derivative of $L_z(t)$ along (26), (50), and (53) is

$$\dot{L}_z(t) = z_1(t)(f_1(\bar{x}_1) - \dot{y}_r(t) + z_2(t) + \hat{\alpha}_1(z_1)) + \sum_{k=2}^{n-1} z_k(t)(f_i(\bar{x}_i) + z_{i+1}(t) - \dot{\hat{\alpha}}_{i-1} + \hat{\alpha}_i(z_i)) + z_n(t)(f_n(\bar{x}_n(t)) - \dot{\hat{\alpha}}_{n-1}(z_{n-1}) + u). \quad (66)$$

Substituting (19), (46), and (60) into (66) has

$$\dot{L}_z(t) = -\beta_1 z_1^2(t) + z_1 f_1(\bar{x}_1) - z_1(t)\dot{y}_r + z_1(t)z_2(t) - \frac{1}{2}z_1(t)\hat{W}_{a1}^T(t)S_1(z_1) + \sum_{k=2}^{n-1}\Big(-\beta_i z_i^2(t) + z_i(t)f_i(\bar{x}_i) - z_i(t)\dot{\hat{\alpha}}_{i-1}(z_{i-1}) + z_i(t)z_{i+1}(t) - \frac{1}{2}z_i\hat{W}_{ai}^T(t)S_i(z_i)\Big) - \beta_n z_n^2 + z_n f_n(\bar{x}_n) - z_n\dot{\hat{\alpha}}_{n-1} - \frac{1}{2}z_n\hat{W}_{an}^T(t)S_n(z_n). \quad (67)$$

Applying Young's inequality $ab \le (a^2/2) + (b^2/2)$ to (67), the following result can be yielded:

$$\dot{L}_z(t) \le -(\beta_1 - 3)z_1^2(t) - \sum_{k=2}^{n}(\beta_i - 4)z_i^2(t) + D(t) \quad (68)$$

where $D(t) = (1/2)\dot{y}_r^2(t) + (1/2)\sum_{k=2}^{n-1}\dot{\hat{\alpha}}_{k-1}^2 + (1/2)\sum_{k=1}^{n} f_k^2(\bar{x}_k) + (1/8)\sum_{k=1}^{n}(\hat{W}_{ak}^T(t)S_k(z_k))^2$. Because it has been proven that $\tilde{W}_{ai}(t)$, $i = 1, \ldots, n$, are SGUUB by part 1, the term $\sum_{k=1}^{n}(\hat{W}_{ak}^T(t)S_k(z_k))^2$ is bounded. Since all terms of $D(t)$ are bounded, there exists a constant ρ such that $|D(t)| < \rho$. Thus, the following result can be held:

$$\dot{L}_z(t) < -\beta L_z(t) + \rho$$

where $\beta = \min\{\beta_1 - 3, \beta_2 - 4, \ldots, \beta_n - 4\}$. Based on the above-obtained result, applying Lemma 1 has

$$L_z(t) < e^{-\beta t}L_z(0) + \frac{\rho}{\beta}(1 - e^{-\beta t}).$$

It implies that the tracking errors can arrive to the desired accuracy by making β large enough, as a result that the desired control performance can be obtained.

$$\xi_n(t) = [z_n(t), \tilde{W}_{an}^T(t), \tilde{W}_{cn}^T(t)]^T$$

$$A_n(t) = \begin{bmatrix} \beta_n - 4 & 0 & 0 \\ 0 & \left(\frac{\gamma_{an}}{2} - \frac{\gamma_{cn}^2}{2} - \frac{1}{32}W_n^{*T}\omega_n\omega_n^T W_n^*\right)S_n(z_n)S_n^T(z_n) & 0 \\ 0 & 0 & \frac{1}{\|\omega_n(t)\|^2+1}\left(\frac{\gamma_{cn}}{2} - \frac{1}{32}W_n^{*T}S_n(z_n)S_n^T(z_n)W_n^*\right)\omega_n(t)\omega_n^T(t) \end{bmatrix}$$

$$C_n(t) = \frac{1}{2}f_n^2(\bar{x}_n) + \frac{1}{2}\dot{\hat{\alpha}}_{n-1}^2 + \frac{\gamma_{an}+1}{2}\left(W_n^{*T}S_n(z_n)\right)^2 + \frac{\gamma_{cn}}{2}\epsilon_n^2(t)$$

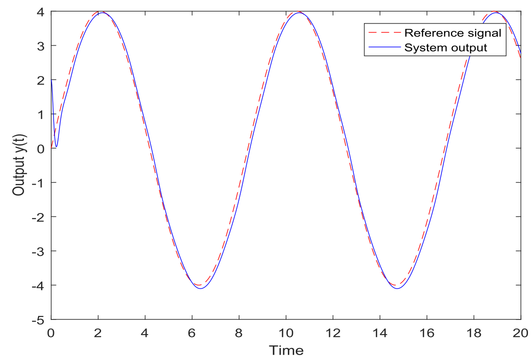


Fig. 1. Tracking performance.

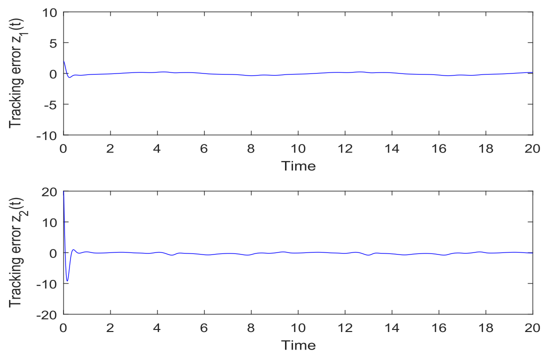


Fig. 2. Tracking errors for the first and second steps.

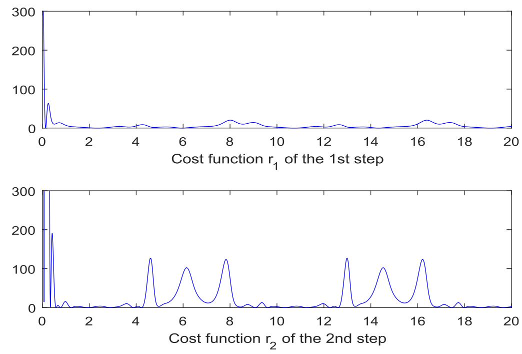


Fig. 3. The cost functions for the first and second steps.

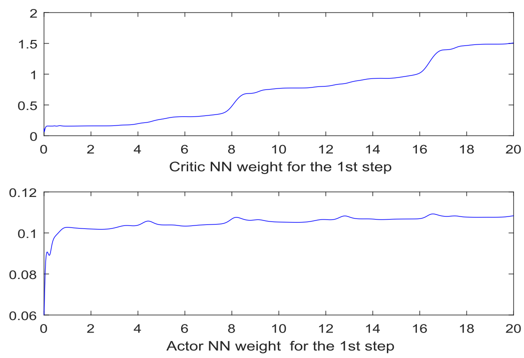


Fig. 4. Critic and actor NN weight norms for the first step.

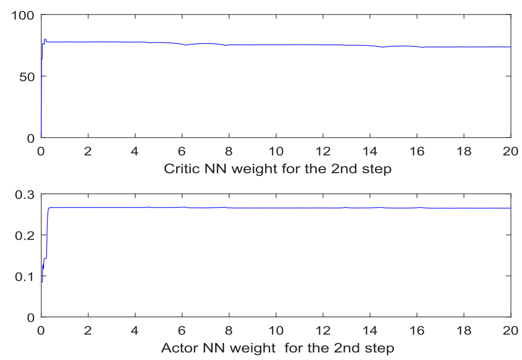


Fig. 5. Critic and actor NN weight norms for the second step.

IV. Simulation Example

In order to further demonstrate the effectiveness of the proposed control technique, a numerical simulation is carried out for a second-order strict-feedback system.

Consider the following nonlinear system:

$$\begin{aligned}\dot{x}_1(t) &= -\sin^2(2x_1) + x_2(t)\\ \dot{x}_2(t) &= (1-(2+\sin(x_1)\cos(x_2))^2) + u \end{aligned}\tag{69}$$

where $x_1(t), x_2(t) \in R$ are the system states and $u \in R$ is the control input. The desired reference signal is $y_r = 4\sin(3t/4)$ shown in Fig. 1.

Step 1: From the system equation (69), the tracking error dynamic for the first backstepping step is $\dot{z}_1(t) = -\sin^2(2x_1) + x_2(t) - 3\cos(3t/4)$. The initial position is $x_1(0) = 2$. The virtual controller is constructed based on (19), and the design parameter is $\beta_1 = 12$.

For the step, the critic and actor NNs contain 36 nodes with centers μ_i evenly spaced in the range $[-6, 6]$, and the widths of the Gaussian function are $\phi_i = 1$, $i = 1, \ldots, 36$. The updating laws for critic and actor NNs are given based on (23) and (24), respectively, of which the learning rates are $\gamma_{c1} = 0.2$ and $\gamma_{a1} = 3$ and the initial conditions are $W_{c1}(0) = [0.02, \ldots, 0.02]^T \in R^{36\times 1}$ and $W_{a1}(0) = [0.01, \ldots, 0.01]^T \in R^{36\times 1}$.

Step 2: This is the final backtepping step, and the actual controller is designed in the step. The error dynamic for the step is $\dot{z}_2(t) = (1-(2+\sin(x_1)\cos(x_2))^2) - \dot{\alpha}_1(t) + u$. The initial position is $x_2(0) = -2$. The actual controller is constructed based on (60), and the design parameter is $\beta_2 = 14$.

For the final step, the critic and actor NNs are constructed to contain 72 nodes with centers μ_i evenly spaced in the range $[-6, 6]$, and the widths of the Gaussian function are $\phi_i = 1$, $i = 1, \ldots, 72$. The updating laws are obtained from (61) and (62), respectively. Their learning rates are $\gamma_{c2} = 0.3$ and $\gamma_{a2} = 4$, and initial conditions are $W_{c2}(0) = [0.02, \ldots, 0.02]^T \in R^{72\times 1}$ and $W_{a2}(0) = [0.01, \ldots, 0.01]^T \in R^{72\times 1}$.

Figs. 1–5 show the simulation results. Fig. 1 shows the tracking performance. Tracking errors $z_1(t)$ and $z_2(t)$ are displayed in Fig. 2, which converge to zero. The cost functions $r_1(z_1, \alpha_1)$ and $r_2(z_2, u)$ are presented in Fig. 3. The boundness of critic and actor weight vectors is shown in Figs. 4 and 5. Figs. 1–5 further demonstrate that the proposed control can guarantee that the control objective is achieved.

In order to demonstrate the optimizing performance of the proposed control method, a comparison with the published control approach proposed in [29] is carried out.

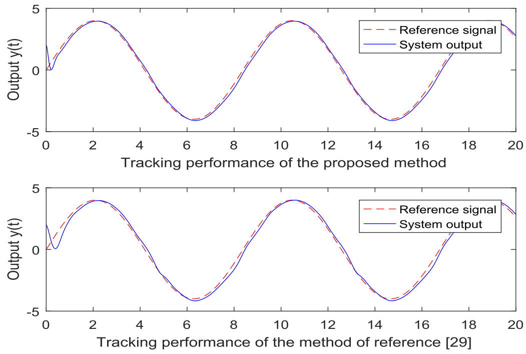


Fig. 6. Two tracking performances.

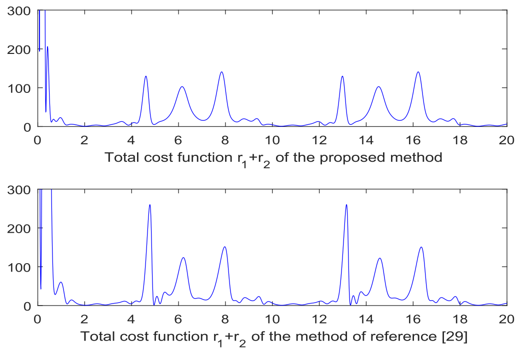


Fig. 7. Two total cost functions.

The comparative results are shown in Figs. 6 and 7. Fig. 6 shows that two tracking performances are the same, and Fig. 7 shows the cost functions of two control schemes. From Figs. 6 and 7, it can be directly concluded that, with the same tracking performances, the proposed control scheme is low cost.

V. Conclusion

This paper proposes a new control technique named OB for strict-feedback systems, which melts the optimization into backstepping control. Since backstepping is the most general and effective control technique for strict-feedback systems, it is very significant and advantageous to consider optimization to the control. In order to achieve the objective, the actor–critic-based RL algorithm is used, in which the actor NN is utilized to carry out the control behavior; the critic NN is utilized to evaluate the optimizing performance and return the evaluation to actor training. Since all the virtual controls and the actual control are designed to be the optimized solutions of corresponding subsystems, the overall control is optimized. Based on the Lyapunov analysis, it is proven that the proposed scheme can achieve the control objective. Simulation results show the effectiveness of the proposed control approach.

Acknowledgment

The authors would like to thank the National Research Foundation, Keppel Corporation, and National University of Singapore for supporting this work done in the Keppel-NUS Corporate Laboratory. The conclusions put forward reflect the views of the authors alone and not necessarily those of the institutions within the Corporate Laboratory. The WBS number of this project is R-261-507-004-281.

References

[1] D. Seto, A. M. Annaswamy, and J. Baillieul, "Adaptive control of nonlinear systems with a triangular structure," *IEEE Trans. Autom. Control*, vol. 39, no. 7, pp. 1411–1428, Jul. 1994.

[2] Z. Pan and T. Basar, "Adaptive controller design for tracking and disturbance attenuation in parametric strict-feedback nonlinear systems," *IEEE Trans. Autom. Control*, vol. 43, no. 8, pp. 1066–1083, Aug. 1998.

[3] G.-X. Wen, Y.-J. Liu, S.-C. Tong, and X.-L. Li, "Adaptive neural output feedback control of nonlinear discrete-time systems," *Nonlinear Dyn.*, vol. 65, nos. 1–2, pp. 65–75, 2011.

[4] G.-X. Wen, Y.-J. Liu, and C. L. P. Chen, "Direct adaptive robust NN control for a class of discrete-time nonlinear strict-feedback SISO systems," *Neural Comput. Appl.*, vol. 21, no. 6, pp. 1423–1431, 2012.

[5] T. Zhang, S. S. Ge, and C. C. Hang, "Adaptive neural network control for strict-feedback nonlinear systems using backstepping design," *Automatica*, vol. 36, no. 12, pp. 1835–1846, 2000.

[6] D. Wang and J. Huang, "Neural network-based adaptive dynamic surface control for a class of uncertain nonlinear systems in strict-feedback form," *IEEE Trans. Neural Netw.*, vol. 16, no. 1, pp. 195–202, Jan. 2005.

[7] C. L. P. Chen, G.-X. Wen, Y.-J. Liu, and Z. Liu, "Observer-based adaptive backstepping consensus tracking control for high-order nonlinear semi-strict-feedback multiagent systems," *IEEE Trans. Cybern.*, vol. 46, no. 7, pp. 1591–1601, Jul. 2016.

[8] W. Li, L. Liu, and G. Feng, "Distributed containment tracking of multiple stochastic nonlinear systems," *Automatica*, vol. 69, pp. 214–221, Jul. 2016.

[9] W. Li, L. Liu, and G. Feng, "Cooperative control of multiple stochastic high-order nonlinear systems," *Automatica*, vol. 82, pp. 218–225, Aug. 2017.

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

[11] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*. Geneva, Switzerland: Interscience Publishers, 1962.

[12] F. L. Lewis, D. Vrabie, and V. L. Syrmos, *Optimal Control*, 3rd ed. Hoboken, NJ, USA: Wiley, 2012.

[13] D. Wang, C. Li, D. Liu, and C. Mu, "Data-based robust optimal control of continuous-time affine nonlinear systems with matched uncertainties," *Inf. Sci.*, vol. 366, pp. 121–133, Oct. 2016.

[14] S. Bhasin, R. Kamalapurkar, M. Johnson, K. 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.

[15] H. Modares, F. L. Lewis, and M.-B. Naghibi-Sistani, "Adaptive optimal control of unknown constrained-input systems using policy iteration and neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, no. 10, pp. 1513–1525, Oct. 2013.

[16] D. Wang, D. Liu, and H. Li, "Policy iteration algorithm for online design of robust control for a class of continuous-time nonlinear systems," *IEEE Trans. Autom. Sci. Eng.*, vol. 11, no. 2, pp. 627–632, Apr. 2014.

[17] D. Wang, D. Liu, Q. Zhang, and D. Zhao, "Data-based adaptive critic designs for nonlinear robust optimal control with uncertain dynamics," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 46, no. 11, pp. 1544–1555, Nov. 2016.

[18] 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.

[19] D. Wang, D. Liu, Q. Wei, D. Zhao, and N. Jin, "Optimal control of unknown nonaffine nonlinear discrete-time systems based on adaptive dynamic programming," *Automatica*, vol. 48, no. 8, pp. 1825–1832, 2012.

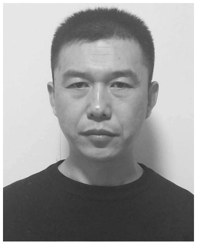
[20] P. J. Werbos, "Approximate dynamic programming for real-time control and neural modeling," *Handbook Intell. Control, Neural, Fuzzy, Adapt. Approaches*, vol. 15, pp. 493–525, Jan. 1992.

[21] 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.

[22] H. Modares, F. L. Lewis, and M.-B. Naghibi-Sistani, "Integral reinforcement learning and experience replay for adaptive optimal control of partially-unknown constrained-input continuous-time systems," *Automatica*, vol. 50, no. 1, pp. 193–202, 2014.

[23] D. Liu, H. Li, and D. Wang, "Online synchronous approximate optimal learning algorithm for multi-player non-zero-sum games with unknown dynamics," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 44, no. 8, pp. 1015–1027, Aug. 2014.

[24] 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, Dec. 1997.
[25] P. He and S. Jagannathan, “Reinforcement learning neural-network-based controller for nonlinear discrete-time systems with input constraints,” *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 37, no. 2, pp. 425–436, Apr. 2007.
[26] S. S. Ge, C. C. Hang, and T. Zhang, “Adaptive neural network control of nonlinear systems by state and output feedback,” *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 29, no. 6, pp. 818–828, Dec. 1999.
[27] W. He, S. S. Ge, Y. Li, E. Chew, and Y. S. Ng, “Impedance control of a rehabilitation robot for interactive training,” in *Proc. Int. Conf. Soc. Robot.*, 2012, pp. 526–535.
[28] G.-X. Wen, C. L. P. Chen, Y.-J. Liu, and Z. Liu, “Neural-network-based adaptive leader-following consensus control for second-order nonlinear multi-agent systems,” *IET Control Theory Appl.*, vol. 9, no. 13, pp. 1927–1934, Aug. 2015.
[29] S. S. Ge and C. Wang, “Direct adaptive NN control of a class of nonlinear systems,” *IEEE Trans. Neural Netw.*, vol. 13, no. 1, pp. 214–221, Jan. 2002.

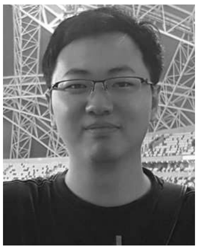


Guoxing Wen received the M.S. degree in applied mathematics from the Liaoning University of Technology, Jinzhou, China, in 2011, and the Ph.D. degree in computer and information science from Macau University, Macau, China, in 2014.

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

Shuzhi Sam Ge (S’90–M’92–SM’99–F’06) received the B.Sc. degree in control engineering from the Beijing University of Aeronautics and Astronautics, Beijing, China, in 1986, and the Ph.D. degree and the Diploma of Imperial College degree in mechanical/electrical engineering from the Imperial College of Science, Technology and Medicine, University of London, London, U.K., in 1993.

He is the Founding Director of the Social Robotics Laboratory, Interactive Digital Media Institute, National University of Singapore, Singapore, where he is currently a Professor with the Department of Electrical and Computer Engineering.



Fangwen Tu received the B.E. degree from the Department of Electrical Engineering, Dalian University of Technology, Dalian, China, in 2012, and the Ph.D. degree from the Department of Electrical and Computer Engineering, National University of Singapore, Singapore, in 2017.

His current research interests include intelligent control, machine learning, data mining, and computer vision.

报告编号：BZU20180032

文献检索证明

作者姓名：文国兴

作者单位：滨州学院

该作者 2018 年发表在期刊《IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS》第 29 卷第 8 期的论文《Optimized Backstepping for Tracking Control of Strict-Feedback Systems》被 SCIE（SCI-EXPANDED）收录。该期刊 2017 年 JCR 的影响因子是 7.982，JCR 分区见附件。在中科院 2016 年 JCR 分区表中位于大类分区 1 区。

检索结果见附件。

证明单位：滨州学院图书馆

证明人：　郭春霞

2018 年 10 月 26 日

关闭 Web of Science 打印
第 1 页 (记录 1 -- 1)
◀[1]▶

第 1 条，共 1 条

标题: Optimized Backstepping for Tracking Control of Strict-Feedback Systems

作者: Wen, GX (Wen, Guoxing); Ge, SS (Ge, Shuzhi Sam); Tu, FW (Tu, Fangwen)

来源出版物: IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS 卷: 29 期: 8 页: 3850-3862 DOI: 10.1109/TNNLS.2018.2803726 出版年: AUG 2018

Web of Science 核心合集中的 "被引频次": 0

被引频次合计: 0

入藏号: WOS:000439627700044

PubMed ID: 29993615

文献类型: Article

地址: [Wen, Guoxing] Binzhou Univ, Coll Sci, Binzhou 256600, Peoples R China.
[Ge, Shuzhi Sam; Tu, Fangwen] Natl Univ Singapore, Dept Elect & Comp Engn, Singapore 117576, Singapore.
[Ge, Shuzhi Sam] Natl Univ Singapore, Interact Digital Media Inst, Social Robot Lab, Singapore 117576, Singapore.
通讯作者地址: Wen, GX (通讯作者)， Binzhou Univ, Coll Sci, Binzhou 256600, Peoples R China.
电子邮件地址: gxwen@live.cn; samge@nus.edu.sg; fangwen_tu@hotmail.com

ISSN: 2162-237X
eISSN: 2162-2388
来源出版物页码计数: 13

关闭 Web of Science 打印
第 1 页 (记录 1 -- 1)
◀[1]▶

IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS

impact factor

7.982 (2017) 7.658 (5 年)

JCR®类别	类别中的排序	JCR分区
COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE	6/132	Q1
COMPUTER SCIENCE, HARDWARE & ARCHITECTURE	2/52	Q1
COMPUTER SCIENCE, THEORY & METHODS	2/103	Q1
ENGINEERING, ELECTRICAL & ELECTRONIC	8/260	Q1

数据来自第 2017 版 Journal Citation Reports

出版商

IEEE-INST ELECTRICAL ELECTRONICS ENGINEERS INC, 445 HOES LANE, PISCATAWAY, NJ 08855-4141 USA

ISSN: 2162-237X
eISSN: 2162-2388

研究领域

Computer Science
Engineering

关闭窗口

Formation Control With Obstacle Avoidance for a Class of Stochastic Multiagent Systems

Guoxing Wen, C. L. Philip Chen, *Fellow, IEEE*, and Yan-Jun Liu

***Abstract*—This paper addresses formation control with obstacle avoidance problem for a class of second-order stochastic nonlinear multiagent systems under directed topology. Different with deterministic multiagent systems, stochastic cases are more practical and challenging because the exogenous disturbances depicted by the Wiener process are considered. In order to achieve control objective, both the leader-follower formation approach and the artificial potential field (APF) method are combined together, where the artificial potential is utilized to solve obstacle avoidance problem. For obtaining good system robustness to the undesired side effects of the artificial potential, H_∞ analysis is implemented. Based on the Lyapunov stability theory, it is proven that control objective can be achieved, of which obstacle avoidance is proven by finding an energy function satisfying that its time derivative is positive. Finally, a numerical simulation is carried out to further demonstrate the effectiveness of the proposed formation schemes.**

***Index Terms*—Directed topology, formation control, obstacle avoidance, stochastic multiagent system, H_∞ analysis.**

I. Introduction

In recent decades, cooperations or coordinations of multiagent systems have received the increasing attention because the research is meeting military and civilian requirements. Their applications can be found in various fields, such as cooperative control of satellite clusters, formation control of unmanned aerial vehicles, distributed optimization of multiple robotic systems, and scheduling of automated highway systems [1]–[4].

Manuscript received July 30, 2017; revised October 10, 2017; accepted November 23, 2017. Date of publication December 20, 2017; date of current version March 6, 2018. This work was supported in part by the Doctoral Scientific Research Staring Fund of Binzhou University under Grant 2016Y14, in part by the National Natural Science Foundation of China under Grant 61572540, in part by the Macau Science and Technology Development Fund (FDCT) under Grant 019/2015/A, Grant 024/2015/AMJ, and Grant 079/2017/A2, and in part by the University of Macau MYR Grants. *(Corresponding author: Guoxing Wen.)*

G. 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 the Dalian Maritime University, Dalian 116026, China, and also with the 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).

Y.-J. Liu is with the College of Science, Liaoning University of Technology, Jinzhou 121001, China (e-mail: liuyanjun@live.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIE.2017.2782229

In the multiagent community, formation control is one of the most fundamental and important research topics, which requires a group of autonomous agents to keep a predefined formation pattern moving in the desired trajectory with velocity. In some sense, it can also be viewed as all autonomous agents to finish a common task by collaboration. Therefore, multiagent formations can be widely applied in the areas of aerospace, industry, entertainment, and other fields. For example, satellite formation can greatly reduce operating costs, improve system stability and reliability, and exceed the ability of multiple single-spacecrafts. In past decades, many formation strategies, such as leader-follower [5], virtual structure [6], and behavior-based [7], have been well developed and applied, where the leader-follower approach is the most popular because of its simplicity and stability.

However, most existing formation control methods are only focused on deterministic multiagent systems, which do not consider any stochastic disturbances. Since information communication in multiagent system control is often interfered by various kinds of stochastic noises, such as thermal noise, channel fading, and quantization effect during encoding and decoding, the stochastic dynamic model can more precisely to describe the practical multiagent engineering than the deterministic case. Although many control techniques developed for deterministic systems have been successfully extended to stochastic dynamic systems, such as backstepping, adaptive observer, reinforcement learning, and nonlinear optimality [8]–[12], these techniques cannot be directly applied to the multiagent control owing to the state coupling problem. Recently, several consensus schemes of stochastic multiagent systems have been reported and received widespread concern [13], [14]. Nevertheless, in comparison with consensus control, formation control is challenging and interesting because the predefined formation configuration is required to maintain.

In the formation control community, the obstacle avoidance problem is still a big challenge because of uncontrollability and complexity [15]–[17]. To solve the problem, artificial potential field (APF) methods are usually considered [18]–[22]. By treating every obstacle as the high-potential point, a repulsive force will be triggered to compel the agent system to bypass the obstacles when any agent moves into a predefined range around obstacles. Furthermore, in order to achieve the ideal control performance, the robustness analysis is necessary to be performed for disturbance environments. Actually, the artificial potential forces will cause undesired side effects after finishing the tasks of obstacle avoidance, so it can be treated as exogenous disturbances. Generally, H_∞ control strategy is first considered for obtaining the desired system robustness when exogenous disturbances enter a system [23]–[26]. However, most existing formation schemes concerning obstacle avoidance are only focused on the deterministic multiagent systems [18]–[22]. In addition, the existing robust control schemes rarely address the multiagent formation [23]–[26]. Furthermore, the H_∞ robust control of stochastic multiagent formation is very difficult and challenging whether control design or stability analysis because Itô differentiation involves not only gradient but also Hessian term (second-order partial derivative term).

Motivated by the above discussion, this paper addresses formation control with obstacle avoidance problems for a class of second order stochastic multiagent systems under directed topologies. The main contributions can be listed as follows.

1) The obstacle avoidance problem of multiagent formation is solved by combining both the artificial potential method and leader-follower formation approach together, which is proven by a novel method.
2) The proposed formation control scheme is developed for stochastic second-order multiagent systems with directed interconnection topology, so it can be applied to a wide class of practical multiagent engineering.
3) H_∞-technique-based robust control is extended to the stochastic multiagent systems.

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

1) R represents real number; R^n denotes real n-dimensional vector space; $R^{n\times m}$ is $n\times m$-dimensional matrix space; I_n is $n\times n$ identity matrix.
2) $\|\cdot\|$ represents 2-norm; E denotes mathematical expectation; $\|\cdot\|_{L_{E_2}} := (E\int_0^\infty \|\cdot\|^2 dt)^{\frac{1}{2}}$.
3) T is the transposition symbol; ∇ is the gradient operator; $\otimes$ denotes Kronecker product.

II. Preliminaries

A. Stochastic System

Consider the following stochastic system:

$$dy(t) = (f(y) + \tau(t))\,dt + g(y)\,d\omega(t) \tag{1}$$

where $y(t)\in R^n$ is the state; $\tau(t)$ is disturbance input; $\omega(t)\in R^r$ is an independent standard Wiener process; $f: R^n \to R^n$, $g: R^n \to R^{n\times r}$ are Lipschitz with $f(0)=0$ and $g(0)=0$.

Definition 1 [27]: For a positive definite, radially unbounded, twice continuously differentiable function $V(y)$ associated with stochastic systems (1), the infinitesimal generator $\mathscr{L}$ is defined as follows:

$$\mathscr{L}[V(y)] = \frac{\partial V}{\partial y}(f(y)+\tau(t)) + \frac{1}{2}\mathrm{Tr}\left\{g^T\frac{\partial^2 V}{\partial y^2}g\right\}. \tag{2}$$

Definition 2 [28]: The equilibrium state $y\equiv 0$ of stochastic system (1) is said to be exponentially mean square stable if there exist k_1 and k_2 such that

$$E\left[\|y(t)\|^2\right] \le k_1\|y(0)\|^2 e^{-k_2 t}. \tag{3}$$

Definition 3 [29]: The H_∞ problem for stochastic system (1) is said to be solved if the following conditions are satisfied:

1) the closed-loop system (1) is exponentially mean-square stable when $\tau(t)=0$;
2) the following inequality is satisfied under zero initial values:

$$\|y(t)\|^2_{L_{E_2}} \le \gamma\|\tau(t)\|^2_{L_{E_2}} \tag{4}$$

where $\gamma>0$ is the noise attenuation level; $\tau(t)\in L_{E_2}([0,\infty);R^{mn})$.

Lemma 1 [9]: Suppose there exist a C^2 positive function $V(t)\in R^n \to R^+$, two constants c_1, c_2, and class K_∞ functions $\nu_1(\cdot), \nu_2(\cdot)$ such that

$$\nu_1(\|y\|) \le V(y) \le \nu_2(\|y\|)$$
$$\mathscr{L}[V(y)] \le -c_1 V(y) + c_2. \tag{5}$$

Then, there is a unique solution of (1) for any initial state $y(0)\in R^n$ almost surely, and the following condition satisfies:

$$E[V(y(t))] \le e^{-c_1 t}V(y(0)) + \left(1-e^{-c_1 t}\right)\frac{c_2}{c_1}. \tag{6}$$

Remark 1: The basic idea of H_∞ control is that the influence of disturbance input $\tau(t)$ on the system output $y(t)$ is attenuated to desired level. Obviously, if zero initial state is satisfied, the H_∞ performance (4) can be rewritten as $\frac{\|y(t)\|^2_{L_{E_2}}}{\|\tau(t)\|^2_{L_{E_2}}} \le \gamma$, which implies that the gain between $y(t)$ and $\tau(t)$ must be equal or less than the prescribed level γ. Therefore, system output can be robust to disturbances by satisfying the H_∞ control performance (4).

B. Algebraic Graph Theory

Let $G=(V,\varepsilon,A)$ denote a directed graph containing n nodes, where $V=\{v_1,v_2,\ldots,v_n\}$, $\varepsilon\subseteq V\times V$, and $A=[a_{ij}]$ are the node set, edge set, and weighted adjacency matrix, respectively. The interconnection topology of multiagent system can be depicted by a graph G, in which every agent is represented by a node. Let $\varepsilon_{ij}=(v_j,v_i)$ be a directed edge, when $\varepsilon_{ij}\in\varepsilon$ if and only if there is the information flowing from agent j to agent i. A directed network G is said to be strongly connected if any two distinct nodes can be connected by a sequence of directed edges. The agent j is said to be a neighbor of agent i if $\varepsilon_{ij}\in\varepsilon$, and all neighbors of agent i are denoted by the set $N_i=\{v_j\in V:\varepsilon_{ij}\in\varepsilon, j\ne i\}$. The adjacency matrix $A=[a_{ij}]$ is used for describing the communication weights among agents, where $a_{ij}>0 \Leftrightarrow \varepsilon_{ij}\in\varepsilon$ and otherwise $a_{ij}=0$ and $a_{ii}=0$. Laplacian matrix of the graph G is defined as

$$L = D - A \tag{7}$$

where $D=\mathrm{diag}\{d_1,d_2,\ldots,d_n\}$, $d_i=\sum_{j=1}^n a_{ij}$. Let $B=\mathrm{diag}\{b_1,b_2,\ldots,b_n\}^T$ denote the communication weights between agents and leader. It is assumed that at least one agent connects with leader, i.e., $b_1+b_2+\cdots+b_n>0$.

C. Artificial Potentials and Virtual Forces

In order to avoid collision with the obstacles, APF methods are employed by taking the obstacles as high potential points, which produce the repulsive forces to expel all agents away from them.

Define the relative position vector $z_{ik}(t)$ between agent i and obstacle o_k as

$$z_{ik}(t) = x_i(t) - o_k, k = 1, \ldots, q \tag{8}$$

where x_i is the position state of agent i. Then, the repulsive potential function is defined as follows.

Definition 4 [30]: The repulsive potential function $P_k(\|z_{ik}(t)\|)$ is a nonnegative differentiable function such that

1) $P_k(\|z_{ik}\|) \to +\infty$ when $\|z_{ik}\| \to \underline{d}_k$, where $\underline{d}_k$ is the minimal separation distance between agents and obstacle k.
2) $P_k(\|z_{ik}\|)$ attains its minimum when $\|z_{ik}\| > \bar{d}_k$, where $\bar{d}_k$ is the distance threshold simulating the repulsion effect, which satisfies $\bar{d}_k > \underline{d}_k$.

The repulsive force is derived from negative gradient of the potential function $P_k(\|z_{ik}\|)$ as

$$p_{ik}(t) = -\nabla_{z_{ik}} P_k(\|z_{ik}\|) = -\nabla_{x_i} P_k(\|z_{ik}\|). \tag{9}$$

By employing the APF method, the possible collisions between agents and obstacles can be avoided. When all agents move away from the obstacles, i.e., $\{x_1, \ldots, x_n\} \notin \Omega_k$ where $\Omega_k = \{x_i \| \|z_{ik}\| \le \bar{d}_k\}$ is a compact set, the repulsive forces arrive the minimum and satisfy $p_{ik}(t) \in L_2[0, T]$. Although the repulsive forces attenuate to the minimum in the situation, they still produce undesired side effects to the control behaviors. In order to ensure formation behaviors to be robust to the undesired side effect, H_∞ analysis is implemented by considering them as the disturbance inputs. When agent i, $i \in \{1, \ldots, n\}$, is moving toward obstacle o_k, $k \in \{1, \ldots, q\}$, i.e., $x_i \in \Omega_k$, the repulsive force $p_{ik}(t)$ will play a role to drive the agent away from the obstacle.

D. Supporting Lemmas

Lemma 2 [31]: A directed graph G is strongly connected if and only if its Laplacian matrix L is irreducible.

Lemma 3 [32]: If the matrix $L = [l_{ij}] \in R^{n\times n}$ satisfy

1) $l_{ij} \le 0, i \ne j, l_{ii} = -\sum_{j=1}^{n} l_{ij}, i = 1, 2, \ldots, n$;
2) L is irreducible.

Then , the following conclusions hold.

1) Real parts of the eigenvalues excepting for the eigenvalue 0 are positive.
2) $[1, 1, \ldots, 1]^T$ is a right eigenvector corresponding to the eigenvalue 0.
3) if $\delta = [\delta_1, \delta_2, \ldots, \delta_n]^T$ is a left eigenvector corresponding to the eigenvalue 0, then its normalization can be chosen so that $\delta_i > 0$ for all $i = 1, 2, \ldots, n$.

Lemma 4 [33]: Let $L = [l_{ij}] \in R^{n\times n}$ be an irreducible matrix such that $l_{ij} = l_{ji} \le 0$ for $i \ne j$, and $l_{ii} = -\sum_{j=1}^{n} l_{ij}$, then all eigenvalues of the matrix

$$\tilde{L} = L + B = \begin{bmatrix} l_{11} + b_1 & \cdots & l_{1n} \\ \vdots & \ddots & \vdots \\ l_{n1} & \cdots & l_{nn} + b_n \end{bmatrix}$$

are positive, where $b_i \ge 0$ satisfies $b_1 + b_2 + \cdots + b_n > 0$.

Lemma 5 (Schur Complement [34]): A linear matrix inequality $\begin{bmatrix} M(x) & P(x) \\ P^T(x) & N(x) \end{bmatrix} > 0$, where $M(x) = M^T(x)$, $N(x) = N^T(x)$, is equivalent to either of the following conditions:

1) $M(x) > 0$, $N(x) - P^T(x)M^{-1}(x)P(x) > 0$;
2) $N(x) > 0$, $M(x) - P(x)N^{-1}(x)P^T(x) > 0$.

Lemma 6: Let $V(t) \in R$ be a positive definite continuous function, if $\mathscr{L}(V(t)) > \beta V(t)$ (or $\mathscr{L}(V(t)) \le \beta V(t)$) is satisfied, then the following inequality holds:

$$E(V(t)) > e^{\beta(t-t_0)} E(V(t_0))$$
$$\left(\text{or } E(V(t)) \le e^{\beta(t-t_0)} E(V(t_0))\right) \tag{10}$$

where $t \ge t_0$, β is a positive constant.

Proof: From $\mathscr{L}(V(t)) > \beta V(t)$ (or $\mathscr{L}(V(t)) \le \beta V(t)$), the following one holds:

$$\frac{d(E(V))}{dt} = E(\mathscr{L}(V)) > \beta E(V)$$
$$\left(\text{or } \frac{d(E(V))}{dt} = E(\mathscr{L}(V)) \le \beta E(V)\right).$$

Further, having

$$\frac{d(E(V))}{E(V)} > \beta dt \left(\text{or } \frac{d(E(V))}{E(V)} \le \beta dt\right).$$

Integrating the above inequality from t to t_0, there is the following one:

$$\ln(E(V))|_{t_0}^{t} > \beta(t - t_0) \ (\text{or } \ln(E(V))|_{t_0}^{t} \le \beta(t - t_0)).$$

The inequality (10) can be obtained by calculating exponent on both sides of the above inequality. ■

III. MAIN RESULTS

A. Problem Formulation and Control Objective

Consider the second-order multiagent systems molded by the following stochastic differential equations:

$$\begin{aligned} dx_i(t) &= v_i(t)dt \\ dv_i(t) &= (f(x_i, v_i) + u_i)\, dt + \phi_i(x_i, v_i)\, dw_i(t) \\ i &= 1, \ldots, n \end{aligned} \tag{11}$$

where $x_i(t) = [x_{i1}(t), \ldots, x_{im}(t)]^T \in R^m$ and $v_i(t) = [v_{i1}(t), \ldots, v_{im}(t)]^T \in R^m$ are the position and velocity states, respectively; $u_i = [u_{i1}, \ldots, u_{im}]^T \in R^m$ is the control input; $f(\cdot) \in R^m$ is the continuously differentiable vector-valued function with $f(0) = 0_m$; $\phi_i(x_i, v_i) \in R$ are the nonzero smooth functions; $w_i(t)$ is the independent m-dimensional standard Wiener process defined on a complete probability space.

Remark 2: For the multiagent dynamic (11), the standard Wiener process $w_i(t)$ is used to represent stochastic disturbances. Since stochastic disturbances inherently exist in almost all physical systems, such as the Gaussian white noise of a communication channel, it is very necessary to research the stochastic case of multiagent systems.

The leader dynamics are described as

$$\dot{x}_r(t) = v_r(t), \dot{v}_r(t) = f(x_r, v_r) \tag{12}$$

where $x_r(t) \in R^m$ and $v_r(t) \in R^m$ are the position and velocity states, respectively.

Assumption 1 [35]: The continuously differentiable vector-valued function $f(\cdot)$ is Lipschitz, i.e., there exist nonnegative constants ρ_{1i}, ρ_{2i} such that

$$\|f(x_i, v_i) - f(x_r, v_r)\| \le \rho_{1i}\|x_i - x_r\| + \rho_{2i}\|v_i - v_r\| \quad i = 1, \ldots, n. \tag{13}$$

Assumption 2 [36]: The smooth function $\phi_i(x_i, v_i)$, $i = 1, \ldots, n$, in differential (11) satisfies the following condition:

$$\phi_i^2(x_i, v_i) \le \zeta_{1i}\|x_i\|^2 + \zeta_{2i}\|v_i\|^2 \tag{14}$$

where ζ_{1i} and ζ_{2i} are two positive constants.

Assumption 3 [37]: The reference signals $x_r(t)$ and $v_r(t)$ are bounded by the constants ϵ_1 and ϵ_2, i.e., $\|x_r\| \le \epsilon_1$, $\|v_r\| \le \epsilon_2$.

Definition 5 (Mean Square Formation [13]): The stochastic multiagent system (11) is said to reach the mean square formation if the following conditions are held for bounded initial condition:

$$\lim_{t\to\infty} E\left(\|x_i(t) - x_r(t) - \eta_i\|^2\right) = 0$$
$$\lim_{t\to\infty} E\left(\|v_i(t) - v_r(t)\|^2\right) = 0, i = 1, \ldots, n \tag{15}$$

where $\eta_i = [\eta_{i1}, \ldots, \eta_{im}]^T$ is a constant vector to denote the predefined relative position between agent i and reference (12).

In this paper, the control objective is to design a H_∞ formation scheme such that the multiagent system (11) satisfies the following conditions:

1) keep the predefined formation pattern in mean square;
2) follow the desired trajectory with velocity in mean square;
3) solve the obstacle avoidance problem in mean square.

In order to achieve the control objective, the error variables between the agents and leader are defined as

$$e_{xi} = x_i(t) - x_r(t) - \eta_i$$
$$e_{vi} = v_i(t) - v_r(t), i = 1, \ldots, n. \tag{16}$$

From (11) and (12), the error dynamics can be derived as

$$de_{xi}(t) = e_{vi}(t)dt, de_{vi}(t) = \left(\tilde{f}_i(t) + u_i\right)dt + \phi_i(x_i, v_i)\, dw_i, i = 1, \ldots, n \tag{17}$$

where $\tilde{f}_i(t) = f(x_i, v_i) - f(x_r, v_r)$.

The (17) is rewritten to the compact form as

$$de(t) = \left(\left(\begin{bmatrix} 0_{n\times n} & I_n \\ 0_{n\times n} & 0_{n\times n} \end{bmatrix} \otimes I_m\right) e(t) + \begin{bmatrix} 0_{nm} \\ \tilde{f}(t) \end{bmatrix} + \begin{bmatrix} 0_{nm} \\ u \end{bmatrix}\right) dt + \left(\begin{bmatrix} 0_{n\times n} \\ \Phi \end{bmatrix} \otimes I_m\right) dw \tag{18}$$

where $e(t) = [e_x^T(t), e_v^T(t)]^T$, $e_x = [e_{x1}^T(t), \ldots, e_{xn}^T(t)]^T$, $e_v = [e_{v1}^T(t), \ldots, e_{vn}^T(t)]^T$, $\tilde{f}(t) = [\tilde{f}_1^T(\cdot), \ldots, \tilde{f}_n^T(\cdot)]^T$, $u = [u_1^T, \ldots, u_n^T]^T$, $\Phi = \text{diag}\{\phi_1, \ldots, \phi_n\}$, and $w = [w_1^T, \ldots, w_n^T]^T$.

B. Formation Control Protocol and Stability Analysis

Define the formation errors with respect to position and velocity as

$$\tilde{e}_{xi}(t) = \sum_{j\in N_i} a_{ij}(x_i - \eta_i - x_j + \eta_j) + b_i(x_i - x_r - \eta_i)$$
$$\tilde{e}_{vi}(t) = \sum_{j\in N_i} a_{ij}(v_i(t) - v_j(t)) + b_i(v_i(t) - v_r(t))$$
$$i = 1, 2, \ldots, n \tag{19}$$

where a_{ij} is the ith row and jth column element of adjacency matrix A; b_i is the connection weight between agent i and leader.

Based on the error variables (16), the terms $\tilde{e}_{vi}(t)$, $\tilde{e}_{vi}(t)$ can be rewritten as

$$\tilde{e}_{xi}(t) = \sum_{j\in N_i} a_{ij}(e_{xi}(t) - e_{xj}(t)) + b_i e_{xi}(t)$$
$$\tilde{e}_{vi}(t) = \sum_{j\in N_i} a_{ij}(e_{vi}(t) - e_{vj}(t)) + b_i e_{vi}(t)$$
$$i = 1, 2, \ldots, n. \tag{20}$$

Design the formation control as

$$u_i = -\alpha_i(\tilde{e}_{xi} + \tilde{e}_{vi}) - \sum_{k=1}^{q} \gamma_{ik} p_{ik}(t), i = 1, 2, \ldots, n \tag{21}$$

where α_i and γ_{ik} are positive design constants and specified later; $p_{ik}(t)$ is the repulsion force defined by the (9).

Substituting (21) into (17), the following result can be obtained:

$$de_{xi}(t) = e_{vi}(t)dt$$
$$de_{vi}(t) = \left(-\alpha_i(\tilde{e}_{xi}(t) + \tilde{e}_{vi}(t)) - \sum_{k=1}^{q} \gamma_{ik} p_{ik}(t) + \tilde{f}_i(t)\right) dt + \phi_i(x_i, v_i)\, dw_i, i = 1, \ldots, n. \tag{22}$$

Transforming (22) to compact form as

$$de(t) = \left(-\left(\begin{bmatrix} 0_{n\times n} & -I_n \\ \Lambda\tilde{L} & \Lambda\tilde{L} \end{bmatrix} \otimes I_m\right) e(t) - \begin{bmatrix} 0_{nm} \\ p(t) \end{bmatrix} + \begin{bmatrix} 0_{nm} \\ \tilde{f}(t) \end{bmatrix}\right) dt + \left(\begin{bmatrix} 0_{n\times n} \\ \Phi \end{bmatrix} \otimes I_m\right) dw \tag{23}$$

where $\Lambda = \text{diag}\{\alpha_1, \ldots, \alpha_n\}$, $p(t) = [(\sum_{k=1}^{q} \gamma_{1k} p_{1k}(z_{1k}))^T, \ldots, (\sum_{k=1}^{q} \gamma_{nk} p_{nk}(z_{nk}))^T]^T$, $\tilde{L} = L + B$, $B = \text{diag}\{b_1, \ldots, b_n\}$.

Theorem 1: Consider the multiagent system (11) with reference signals (12) under strongly connected communication graph G. The H_∞ formation control (21) can achieve the control objective for bounded initial condition if the design parameters

α_i, γ_{ik}, and κ satisfy the following conditions:

$$\alpha_i = \kappa\delta_i, \gamma_{ik} > 1, i = 1, 2, \dots, n,$$
$$\kappa \geq \frac{\max_{1\leq i\leq n}\{4\rho_{1i} + 3\rho_{2i} + m(\zeta_{1i} + \zeta_{2i})\} + 3}{\lambda_{\min}(\Theta + 2\Delta B)} \tag{24}$$

where $\delta = [\delta_1, \delta_2, \dots, \delta_n]^T$ is the normalized left eigenvector of Laplacian matrix L associated with eigenvalue 0, $\lambda_{\min}(\Theta + 2\Delta B)$ is the minimum eigenvalue of symmetrical matrix $\Theta + 2\Delta B$, $\Theta = L^T\Delta + \Delta L$, $\Delta = \text{diag}\{\delta_1, \delta_2, \dots, \delta_n\}$.

Remark 3: The proof is consisted of two parts, in which part 1 proves the formation performance and part 2 proves the obstacle avoidance. When all agents are not in the area of possible collision, i.e., $\{x_1, \dots, x_n\} \notin \bigcup_{k=1}^q \Omega_k$, although the repulsive force term $\sum_{k=1}^q \gamma_{ik} p_{ik}(z_{ik})$ attains to the minimum, they still affect the formation behavior. In order to obtain the desired robustness, H_∞ analysis is implemented by handling them as disturbance inputs. When any agent enters the scope of possible collision, i.e., $\forall x_i \in \bigcup_{k=1}^q \Omega_k$, the repulsive force term $\sum_{k=1}^q \gamma_{ik} p_{ik}(z_{ik})$ will dramatically increase to drive the agent away from the obstacles.

Proof:

1) Part 1: Choose the following Lyapunov function candidate:

$$V(t) = \frac{1}{2}e^T(t)(Q \otimes I_m)e(t) \tag{25}$$

where $Q = \begin{bmatrix} \kappa(\Theta+2\Delta B) & I_n \\ I_n & I_n \end{bmatrix}$. It should be mentioned that the matrix Q can be reexpressed as $Q = \begin{bmatrix} \tilde{L}^T\Lambda + \Lambda\tilde{L} & I_n \\ I_n & I_n \end{bmatrix}$ by using these facts $\alpha_i = \kappa\delta_i, i = 1, \dots, n$, of condition (24).

According to Lemma 3, the left eigenvector $\delta = [\delta_1, \delta_2, \dots, \delta_n]^T$ of Laplacian matrix L satisfies $\delta_i > 0$, $i = 1, 2, \dots, n$. From the fact $\Theta\mathbf{1}_n = (L^T\Delta + \Delta L)\mathbf{1}_n = L^T\Delta\mathbf{1}_n + \Delta L\mathbf{1}_n = L^T\delta + \Delta L\mathbf{1}_n = 0$, it can be concluded that Θ is a zero row-sum matrix. According to Lemmas 2 and 4, $\Theta + 2\Delta B$ is a positive definite matrix, thus, $\kappa(\Theta + 2\Delta B) - I_n > 0$ can be held if κ satisfies the condition (24). Therefore, the matrix Q is positive definite in accordance with Lemma 5.

The infinitesimal generator of $V(t)$ associating with error dynamic (23) is

$$\mathscr{L}(V(t)) = -\frac{1}{2}e^T(t)\left(\left(\begin{bmatrix} 0_{n\times n} & -I_n \\ \Lambda\tilde{L} & \Lambda\tilde{L}\end{bmatrix}^T Q + Q\begin{bmatrix} 0_{n\times n} & -I_n \\ \Lambda\tilde{L} & \Lambda\tilde{L}\end{bmatrix}\right)\otimes I_m\right)e(t) - e^T(t)(Q\otimes I_m)\left(\begin{bmatrix}0_{nm} \\ p(t)\end{bmatrix} - \begin{bmatrix}0_{nm} \\ \tilde{f}(t)\end{bmatrix}\right) + \frac{1}{2}Tr\left(\left(\begin{bmatrix}0_{n\times n} \\ \Phi\end{bmatrix}^T Q \begin{bmatrix}0_{n\times n} \\ \Phi\end{bmatrix}\right)\otimes I_m\right). \tag{26}$$

Applying to matrix theory, there is the following result:

$$\begin{bmatrix} 0_{n\times n} & -I_n \\ \Lambda\tilde{L} & \Lambda\tilde{L}\end{bmatrix}^T Q + Q\begin{bmatrix} 0_{n\times n} & -I_n \\ \Lambda\tilde{L} & \Lambda\tilde{L}\end{bmatrix} = \begin{bmatrix}\kappa(\Theta+2\Delta B) & 0_{n\times n} \\ 0_{n\times n} & \kappa(\Theta+2\Delta B) - 2I_n\end{bmatrix}. \tag{27}$$

Substituting (27) into (26) yields

$$\mathscr{L}(V(t)) = -\frac{1}{2}e^T(t)\times\left(\begin{bmatrix}\kappa(\Theta+2\Delta B) & 0_{n\times n} \\ 0_{n\times n} & \kappa(\Theta+2\Delta B) - 2I_n\end{bmatrix}\otimes I_m\right)e(t) - e^T(t)(Q\otimes I_m)\left(\begin{bmatrix}0_{nm} \\ p(t)\end{bmatrix} - \begin{bmatrix}0_{nm} \\ \tilde{f}(t)\end{bmatrix}\right) + \frac{m}{2}Tr\left(\Phi^T\Phi\right). \tag{28}$$

Using the following fact

$$e^T(t)(Q\otimes I_m)\left(\begin{bmatrix}0_{nm} \\ p(t)\end{bmatrix} - \begin{bmatrix}0_{nm} \\ \tilde{f}(t)\end{bmatrix}\right) = e^T(t)\begin{bmatrix}p(t) \\ p(t)\end{bmatrix} - \left(e_x^T(t) + e_v^T(t)\right)\tilde{f}(t) \tag{29}$$

the inequality (28) can become

$$\mathscr{L}(V(t)) = -\frac{1}{2}e^T(t)\times\left(\begin{bmatrix}\kappa(\Theta+2\Delta B) & 0_{n\times n} \\ 0_{n\times n} & \kappa(\Theta+2\Delta B) - 2I_n\end{bmatrix}\otimes I_m\right)e(t) - e^T(t)\begin{bmatrix}p(t) \\ p(t)\end{bmatrix} + \left(e_x^T(t) + e_v^T(t)\right)\tilde{f}(t) + \frac{m}{2}\sum_{i=1}^n \phi_i^2. \tag{30}$$

Based on Assumptions 1–3, the following results can be obtained by using Cauchy–Buniakowsky–Schwarz inequality, $(\sum_{k=1}^n a_k b_k)^2 \leq \sum_{k=1}^n a_k^2 \sum_{k=1}^n b_k^2$, and Young's inequality, $ab \leq \frac{a^2}{2} + \frac{b^2}{2}$:

$$\begin{aligned} e_x^T(t)\tilde{f}(t) &\leq \sum_{i=1}^n \left(\|e_{xi}\|\,\|(f(x_i, v_i) - f(x_r, v_r))\|\right) \\ &\leq \sum_{i=1}^n \left(\|e_{xi}\|\left(\rho_{1i}\|e_{xi}\| + \rho_{2i}\|e_{vi}\| + \rho_{1i}\|\eta_i\|\right)\right) \\ &\leq \sum_{i=1}^n \left(\frac{3\rho_{1i} + \rho_{2i}}{2}\|e_{xi}\|^2 + \frac{\rho_{2i}}{2}\|e_{vi}\|^2 + \frac{\rho_{1i}}{2}\|\eta_i\|^2\right) \end{aligned} \tag{31}$$

$$\begin{aligned} e_v^T(t)\tilde{f}(t) &\leq \sum_{i=1}^n \left(\|e_{vi}\|\,\|(f(x_i, v_i) - f(x_r, v_r))\|\right) \\ &\leq \sum_{i=1}^n \left(\|e_{vi}\|\left(\rho_{1i}\|e_{xi}\| + \rho_{2i}\|e_{vi}\| + \rho_{1i}\|\eta_i\|\right)\right) \\ &\leq \sum_{i=1}^n \left(\frac{\rho_{1i}}{2}\|e_{xi}\|^2 + (\rho_{1i} + \rho_{2i})\|e_{vi}\|^2 + \frac{\rho_{1i}}{2}\|\eta_i\|^2\right) \end{aligned} \tag{32}$$

$$\sum_{i=1}^{n}\phi_i^2(x_i,v_i)\le\sum_{i=1}^{n}\left(\zeta_{1i}\|x_i\|^2+\zeta_{2i}\|v_i\|^2\right)$$
$$\le 2\sum_{i=1}^{n}\Big(\zeta_{1i}\|e_{xi}\|^2+\zeta_{2i}\|e_{vi}\|^2+2\zeta_{1i}\|x_r\|^2$$
$$+\zeta_{2i}\|v_r\|^2+2\zeta_{1i}\|\eta_i\|^2\Big)$$
$$\le 2\sum_{i=1}^{n}\left(\zeta_{1i}\|e_{xi}\|^2+\zeta_{2i}\|e_{vi}\|^2\right)+2\sum_{i=1}^{n}\Big(2\zeta_{1i}\epsilon_1^2$$
$$+\zeta_{2i}\epsilon_2^2+2\zeta_{1i}\|\eta_i\|^2\Big). \tag{33}$$

Substituting the above inequalities into (30), the following one can be yielded:

$$\mathscr{L}(V(t))\le-\frac{1}{2}e^T(t)$$
$$\times\left(\begin{bmatrix}\kappa(\Theta+2\Delta B)-N_1 & 0_{n\times n}\\ 0_{n\times n} & \kappa(\Theta+2\Delta B)-N_2-2I_n\end{bmatrix}\right.$$
$$\left.\otimes I_m\right)e(t)-e^T(t)\begin{bmatrix}p(t)\\p(t)\end{bmatrix}$$
$$+\sum_{i=1}^{n}\left(2m\zeta_{1i}\epsilon_1^2+m\zeta_{2i}\epsilon_2^2+(2m\zeta_{1i}+\rho_{1i})\|\eta_i\|^2\right) \tag{34}$$

where

$$N_1=\begin{bmatrix}4\rho_{11}+\rho_{21}+2m\zeta_{11} & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & 4\rho_{1n}+\rho_{2n}+2m\zeta_{1n}\end{bmatrix},$$
$$N_2=\begin{bmatrix}2\rho_{11}+3\rho_{21}+2m\zeta_{21} & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & 2\rho_{1n}+3\rho_{2n}+2m\zeta_{2n}\end{bmatrix}.$$

Adding and subtracting the term $\frac{1}{4}[p^T(t),p^T(t)][p^T(t),p^T(t)]^T$ on the right-hand side of inequality (34), the following inequality is yielded:

$$\mathscr{L}(V(t))\le-\frac{1}{2}e^T(t)$$
$$\times\left(\begin{bmatrix}\kappa(\Theta+2\Delta B)-N_1-I_n & 0_{n\times n}\\ 0_{n\times n} & \kappa(\Theta+2\Delta B)-N_2-3I_n\end{bmatrix}\right.$$
$$\left.\otimes I_m\right)e(t)-\Bigg(e(t)$$
$$+\frac{1}{2}\begin{bmatrix}p(t)\\p(t)\end{bmatrix}\Bigg)^T\left(e(t)+\frac{1}{2}\begin{bmatrix}p(t)\\p(t)\end{bmatrix}\right)+\frac{1}{4}\left\|\begin{bmatrix}p(t)\\p(t)\end{bmatrix}\right\|^2$$
$$+\sum_{i=1}^{n}\left(2m\zeta_{1i}\epsilon_1^2+m\zeta_{2i}\epsilon_2^2+(2m\zeta_{1i}+\rho_{1i})\|\eta_i\|^2\right). \tag{35}$$

From the fact that $(e+\frac{1}{2}[\begin{smallmatrix}p(t)\\p(t)\end{smallmatrix}])^T(e+\frac{1}{2}[\begin{smallmatrix}p(t)\\p(t)\end{smallmatrix}])\ge 0$, (35) is rewritten as

$$\mathscr{L}(V(t))\le-\frac{1}{2}e^T(t)(M\otimes I_m)e(t)+\gamma\xi^T(t)\xi(t) \tag{36}$$

where

$$M=\begin{bmatrix}\kappa(\Theta+2\Delta B)-N_1-I_n & 0_{n\times n}\\ 0_{n\times n} & \kappa(\Theta+2\Delta B)-N_2-3I_n\end{bmatrix},$$
$$\gamma=1/2\max_{i=1,\ldots,n}\left\{\sum_{k=1}^{p}\gamma_{ik}^2\right\},$$
$$\xi(t)=\Bigg[\sqrt{\frac{\sum_{i=1}^{n}\left(2m\zeta_{1i}\epsilon_1^2+m\zeta_{2i}\epsilon_2^2+(2m\zeta_{1i}+\rho_{1i})\|\eta_i\|^2\right)}{\gamma}},$$
$$\left(\sum_{k=1}^{p}p_{1k}(z_{1k})\right)^T,\ldots,\left(\sum_{k=1}^{p}p_{nk}(z_{nk})\right)^T\Bigg]^T.$$

Since M is a positive definite matrix when designing κ satisfies (24), the inequality (36) can be rewritten as

$$\mathscr{L}(V(t))\le-\frac{\lambda_{\min}(M)}{\lambda_{\max}(Q)}V(t)+\gamma\xi^T(t)\xi(t). \tag{37}$$

The repulsive force $p(t)$ is handled as the disturbance input in the case. If $p(t)=0$, the inequality (37) can be rewritten as

$$\mathscr{L}(V(t))\le-c_1V(t)+c_2 \tag{38}$$

where $c_1=\frac{\lambda_{\min}(M)}{2\lambda_{\max}(Q)}$, $c_2=\sum_{i=1}^{n}(2m\zeta_{1i}\epsilon_1^2+m\zeta_{2i}\epsilon_2^2+(2m\zeta_{1i}+\rho_{1i})\|\eta_i\|^2)$.

From Lemma 1, the following one can be obtained:

$$E[V(t)]\le e^{-c_1t}V(0)+(1-e^{-c_1t})\tfrac{c_2}{c_1}. \tag{39}$$

By making the design parameter κ large enough, the formation errors convergence to desired accuracy, which implies the exponentially mean square stable to be achieved.

Since the multiagent systems get far from the obstacles, $\xi(t)$ belongs to $L_{E_2}([0,\infty);R^{mn})$. By Integrating (36) from 0 to T and taking expectation, the following results can be obtained

$$E\int_0^T\mathscr{L}(V(t))\,dt=E(V(T)-V(0))$$
$$\le-\frac{\lambda_{\min}(M)}{2}E\int_0^T\|e(t)\|^2dt+\gamma E\int_0^T\|\xi(t)\|^2dt$$
$$=-\beta\|e(t)\|_{L_{E_2}}^2+\gamma\|\xi(t)\|_{L_{E_2}}^2 \tag{40}$$

where $\beta=\frac{\lambda_{\min}(M)}{2}$.

Obviously, $\|e(t)\|_{L_{E_2}}^2\le\frac{\gamma}{\beta}\|\xi(t)\|_{L_{E_2}}^2$ if $V(0)=0$, and thus, the H_∞ control performance (4) is satisfied.

2) Part 2: (In the part, collision avoidance is analyzed only for agent i and obstacle j. For the other cases, the proofs are similar.)

Consider the following energy function:

$$V_{ij}(t)=\frac{1}{2}z_{ij}^T(t)z_{ij}(t)+\frac{1}{2}v_i^T(t)v_i(t). \tag{41}$$

Using (11), the infinitesimal generator is

$$\mathscr{L}\left(V_{ij}(t)\right)=z_{ij}^{T}v_i-\alpha_i {v_i}^{T}\left(\tilde{e}_{xi}(t)+\tilde{e}_{vi}(t)\right)+{v_i}^{T}\tilde{f}_i(t) - {v_i}^{T}\sum_{k=1,k\neq j}^{p}\gamma_{ik}p_{ik}(t)-\gamma_{ij}{v_i}^{T}p_{ij}(t)+\frac{1}{2}\phi_i^2. \quad (42)$$

Since the dwell time of agent i in the region Ω_j is finite, these continuous terms $z_{ij}(t)$, $v_i(t)$, $\tilde{e}_{xi}(t)$, $\tilde{e}_{vi}(t)$, $\tilde{f}_i(t)$, ϕ_i and $\sum_{k=1,k\neq j}^{p}\gamma_{ik}p_{ik}(t)$ are bounded. In addition, if the agent i is closing the obstacle j, it implies that the agent is moving toward gradient direction of the artificial potential $P_j(t)$, from the definition of repulsive potential (Definition 2), there is the fact that $-v_i^T(t)p_{ij}(t)=v_i^T\nabla_{x_i}P_j(t)\rightarrow\infty$ if $\|z_{ij}\|\rightarrow \underline{d}_j$. Therefore, the following inequality can be held if agent i is closing to obstacle j sufficiently:

$$-\gamma_{ij}v_i^T(t)p_{ij}(t)>\frac{\gamma_{ij}}{2}z_{ij}^{T}z_{ij}+\frac{\gamma_{ij}}{2}v_i^{T}v_i-z_{ij}^{T}v_i-{v_i}^{T}\tilde{f}_i(t) +\alpha_i{v_i}^{T}\left(\tilde{e}_{xi}(t)-\tilde{e}_{vi}(t)\right)+\frac{1}{2}\phi_i^2+\sum_{k=1,k\neq j}^{p}\gamma_{ik}p_{ik}(t). \quad (43)$$

Applying the above fact to (42) yields

$$\mathscr{L}\left(V_{ij}(t)\right)>\gamma_{ij}V_{ij}(t). \quad (44)$$

According to Lemma 6, the following result holds:

$$E(\|z_{ij}(t)\|^2)>2e^{\gamma_{ij}(t-t_0)}E\left(V_{ij}(t_0)\right)-E(\|v_i(t)\|^2). \quad (45)$$

Thus, $\|z_{ij}(t)\|>\underline{d}_j$ can be guaranteed by designing the parameter γ_{ij} appropriately, i.e., the collision between agent i and obstacle j can be avoided in mean square. ■

IV. Simulation Examples

In order to demonstrate the effectiveness of the proposed control strategy, a simulation example of stochastic multiagent formation that is consisted of four agents is carried out. The multiagent system is modeled as

$$dx_i(t)=v_i(t)dt$$
$$dv_i(t)=\left(\begin{bmatrix}5\cos\left(0.1v_{i1}\right)\\3\sin\left(0.2v_{i2}\right)\end{bmatrix}+u_i\right)dt+\frac{\|v_i\|}{\|x_i\|}dw_i(t)$$
$$i=1,\ldots,4. \quad (46)$$

Their initial positions are $x_1(0)=[6,5]$, $x_2(0)=[-5,6]$, $x_3(0)=[5,-6]$, $x_4(0)=[-6,-5]$.

The reference signal is modeled by the following dynamic:

$$\dot{x}_i(t)=v_i(t),\dot{v}_i(t)=\begin{bmatrix}5\cos\left(0.1x_{i1}(t)\right)\\3\sin\left(0.2x_{i2}(t)\right)\end{bmatrix}. \quad (47)$$

The desired formation pattern is $\eta_1=[4;4]$, $\eta_2=[-4;4]$, $\eta_3=[4;-4]$, $\eta_4=[-4;-4]$. Two obstacle points, o_1 and o_2 , are set at $t=4.2$ and $t=14$, respectively. The desired trajectory and two obstacles are presented in Fig. 1.

Control objective: by applying the control protocol (21), steering the multiagent system (46) follows to the reference signals (47), meanwhile maintains the predefined formation pattern and avoids collision with obstacles.

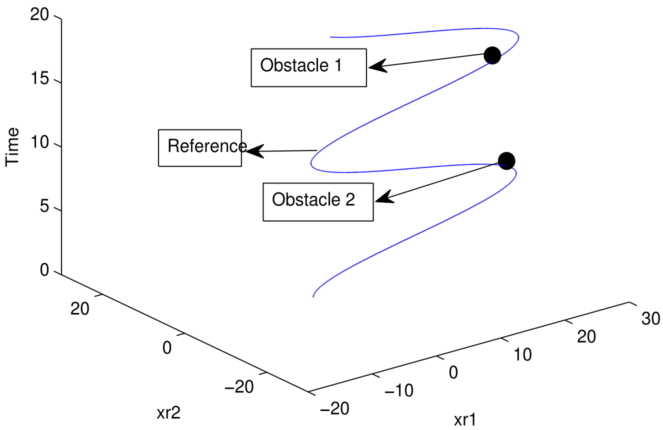


Fig. 1. Reference trajectory with two obstacles.

The Laplacian matrix is

$$L=\begin{bmatrix}1.5&-0.7&0&-0.8\\-0.6&1.4&-0.8&0\\-0.8&0&1.7&-0.9\\0&-0.7&-0.9&1.6\end{bmatrix}.$$

The weight matrix between agents and leader is $B=\operatorname{diag}\{0,0.9,0,0.9\}$.

The potential functions are designed as

$$P_1\left(\|z_{i1}\|\right)=\|z_{i1}(t)\|\,e^{(\|z_{i1}(t)\|-5)^{-2}}$$
$$P_2\left(\|z_{i2}\|\right)=\|z_{i2}(t)\|\,e^{(\|z_{i2}(t)\|-4)^{-2}}. \quad (48)$$

The corresponding repulsive forces derived from negative gradient of the potential functions are

$$p_{i1}=-\nabla_{x_i}P_1\left(\|z_{i1}\|\right)=\Big(2(\|z_{i1}\|-5)^{-3}e^{(\|z_{i1}\|-5)^{-2}} -\|z_{i1}\|^{-1}\,e^{(\|z_{i1}\|-5)^{-2}}\Big)z_{i1}(t)$$
$$p_{i2}=-\nabla_{x_i}P_2\left(\|z_{i2}\|\right)=\Big(2(\|z_{i2}\|-4)^{-3}e^{(\|z_{i2}\|-4)^{-2}} -\|z_{i2}\|^{-1}\,e^{(\|z_{i2}\|-4)^{-2}}\Big)z_{i2}(t),i=1,\ldots,4. \quad (49)$$

The simulation results are shown in Figs. 2–5. Fig. 2 displays the formation control without the assistance of artificial potentials, where the controller is $u_i=-50\left(\tilde{e}_{xi}(t)+\tilde{e}_{vi}(t)\right), i=1,2,3,4$. Obviously, the obstacle avoidance cannot be achieved. In order to solve the problem, the artificial potentials (49) are employed in accordance with the formation protocol (21). Then, the controller is derived as $u_i=-50\left(\tilde{e}_{xi}+\tilde{e}_{vi}\right)-41.5p_{i1}(t)-36p_{i2}(t), i=1,2,3,4$, where 50, 41.5, and 36 are the controller parameters α_i, γ_{i1}, and γ_{i2} of (49), respectively. Fig. 3 shows the control performance under the assistance of artificial potentials. Obviously, the obstacle avoidance can be achieved. Fig. 4 shows the velocity error of multiagent formation, it implies that all agents can follow the reference velocity after finishing the obstacle avoidance. The comparison between both with and without artificial potential is shown in Fig. 5. The simulation

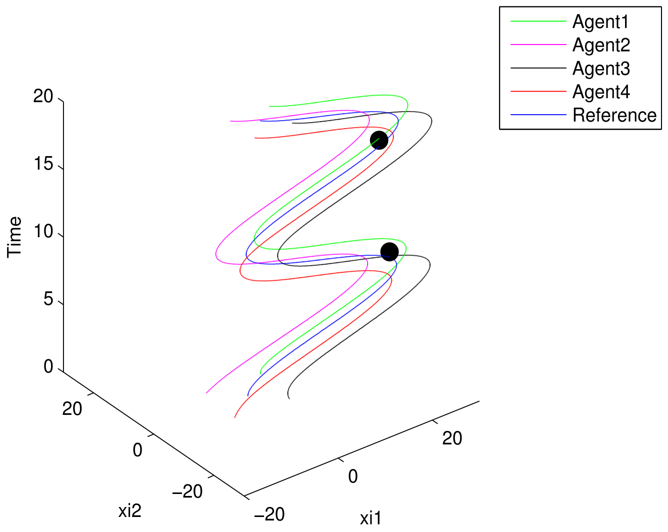


Fig. 2. Obstacle avoidance cannot be achieved without the assistance of artificial potentials.

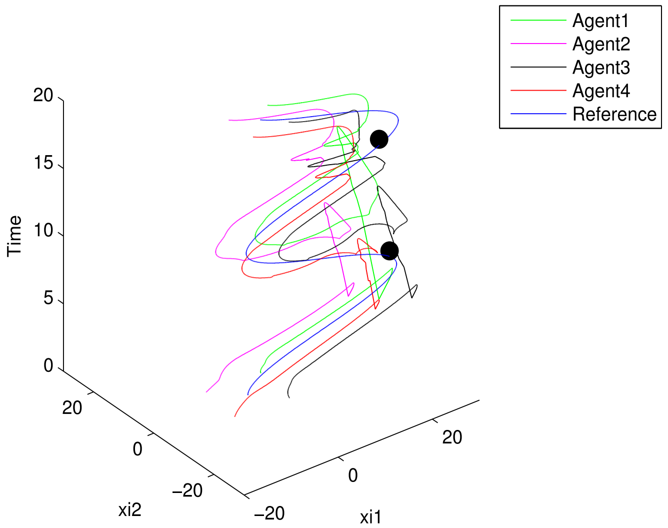


Fig. 3. Obstacle avoidance is achieved under the assistance of artificial potentials.

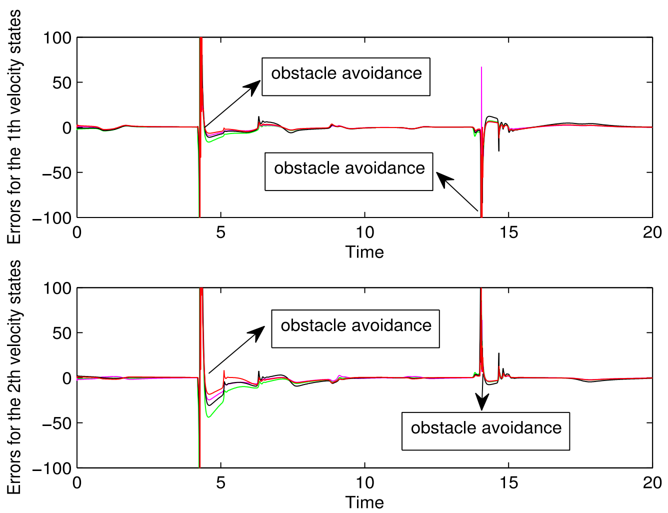


Fig. 4. Velocity errors in the obstacle environment.

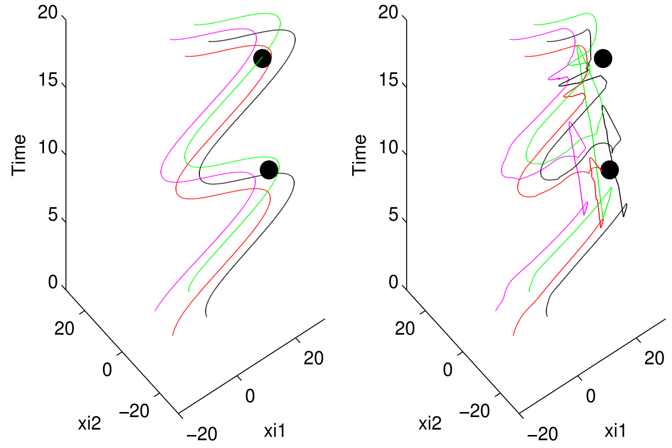


Fig. 5. Comparison concerning obstacle avoidance performance between both with and without the assistance of artificial potential.

results further demonstrate that the proposed stochastic formation approach can well solve the obstacle avoidance problem.

V. Conclusion

The H_∞-technique-based formation control scheme was proposed for second-order stochastic multiagent systems under directed topology. In order to solve the obstacle avoidance problem, APF methods were employed to drive all agents away from obstacles. According to Lyapunov stability theory, it was proven that the proposed formation approach can guarantee the multiagent systems will move along the desired route with velocity while maintaining the predefined formation patterns and avoiding collision with obstacles. Finally, a numerical simulation was carried out to verify the effectiveness of the proposed approach.

References

[1] W. Kang and H.-H. Yeh, "Coordinated attitude control of multi-satellite systems," *Int. J. Robust Nonlinear Control*, vol. 12, no. 2/3, pp. 185–205, 2002.

[2] A. Abdessameud and A. Tayebi, "Formation control of VTOL unmanned aerial vehicles with communication delays," *Automatica*, vol. 47, no. 11, pp. 2383–2394, 2011.

[3] H. Yamaguchi, T. Arai, and G. Beni, "A distributed control scheme for multiple robotic vehicles to make group formations," *Robot. Auton. Syst.*, vol. 36, no. 4, pp. 125–147, 2001.

[4] H. Raza and P. Ioannou, "Vehicle following control design for automated highway systems," *IEEE Control Syst.*, vol. 16, no. 6, pp. 43–60, Dec. 1996.

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

[6] 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.

[7] T. Balch and R. C. Arkin, "Behavior-based formation control for multi-robot teams," *IEEE Trans. Robot. Autom.*, vol. 14, no. 6, pp. 926–939, Dec. 1998.

[8] S.-J. Liu, J.-F. Zhang, and Z.-P. Jiang, "Decentralized adaptive output-feedback stabilization for large-scale stochastic nonlinear systems," *Automatica*, vol. 43, no. 2, pp. 238–251, 2007.

[9] S. Tong, Y. Li, Y. Li, and Y. Liu, "Observer-based adaptive fuzzy backstepping control for a class of stochastic nonlinear strict-feedback systems," *IEEE Trans. Syst., Man, Cybern., B, Cybern.*, vol. 41, no. 6, pp. 1693–1704, Dec. 2011.

[10] R. Cui, L. Chen, C. Yang, and M. Chen, "Extended state observer-based integral sliding mode control for an underwater robot with unknown disturbances and uncertain nonlinearities," *IEEE Trans. Ind. Electron.*, vol. 64, no. 8, pp. 6785–6795, Aug. 2017.

[11] H.-B. Ji and H.-S. Xi, "Adaptive output-feedback tracking of stochastic nonlinear systems," *IEEE Trans. Autom. Control.*, vol. 51, no. 2, pp. 355–360, Feb. 2006.

[12] R. Cui, C. Yang, Y. Li, and S. Sharma, "Adaptive neural network control of AUVS with control input nonlinearities using reinforcement learning," *IEEE Trans. Syst. Man Cybern., Syst.*, vol. 47, no. 6, pp. 1019–1029, Jun. 2017.

[13] T. Li and J.-F. Zhang, "Consensus conditions of multi-agent systems with time-varying topologies and stochastic communication noises," *IEEE Trans. Autom. Control*, vol. 55, no. 9, pp. 2043–2057, Sep. 2010.

[14] Y. Zheng, W. Chen, and L. Wang, "Finite-time consensus for stochastic multi-agent systems," *Int. J. Control*, vol. 84, no. 10, pp. 1644–1652, 2011.

[15] J. Yan, X. Guan, X. Luo, and C. Chen, "Formation control and obstacle avoidance for multi-agent systems based on virtual leader-follower strategy," *Int. J. Inf. Technol. Decision Making*, vol. 23, no. 1, pp. 1–16, 2013.

[16] Y. Liu and J. Hu, "Obstacle avoidance with formation kept for multi-agent," *Int. J. Control Autom.*, vol. 8, pp. 67–74, 2015.

[17] N. Niizuma, K. Nakano, T. Funato, and K. Higuchi, "Group motion control of multi-agent systems with obstacle avoidance: Column formation under input constraints," *Artif. Life Robot.*, vol. 20, no. 1, pp. 70–77, 2015.

[18] S. Cristaldi, A. Ferro, R. Giugno, G. Pigola, and A. Pulvirenti, "Obstacles constrained group mobility models in event-driven wireless networks with movable base stations," *Ad Hoc Netw.*, vol. 9, no. 3, pp. 400–417, 2011.

[19] J. Y. Dai, L. F. Yin, C. Peng, B. J. Yang, and C. S. Wang, "Research on multi-agent formation's obstacle avoidance problem based on three-dimensional vectorial artificial potential field method," *Appl. Mech. Mater.*, vol. 596, no. 5, pp. 251–258, 2014.

[20] Z. Ying and L. Xu, "Leader-follower formation control and obstacle avoidance of multi-robot based on artificial potential field," in *Proc. Control Decision Conf.*, 2015, pp. 4355–4360.

[21] B. Yongqiang, X. Wenzhi, F. Hao, L. Jixiang, and C. Jing, "Obstacle avoidance for multi-agent systems based on stream function and hierarchical associations," in *Proc. IEEE 2012 31st Chin. Control Conf.*, 2012, pp. 6363–6367.

[22] Q. Wang, H. Fang, J. Chen, Y. Mao, and L. Dou, "Flocking with obstacle avoidance and connectivity maintenance in multi-agent systems," in *Proc. 2012 IEEE 51st Annu. Conf. Decision Control*, 2012, pp. 4009–4014.

[23] A. Isidori and A. Astolfi, "Disturbance attenuation and H_∞ control via measurement feedback in nonlinear systems," *IEEE Trans. Autom. Control*, vol. 37, no. 9, pp. 1283–1293, Sep. 1992.

[24] W. M. McEneaney, "Robust H_∞ filtering for nonlinear systems," *Syst. Control Lett.*, vol. 33, no. 5, pp. 315–325, 1998.

[25] C.-S. Tseng and B.-S. Chen, "H_∞ fuzzy estimation for a class of nonlinear discrete-time dynamic systems," *IEEE Trans. Signal Process.*, vol. 49, no. 11, pp. 2605–2619, Nov. 2001.

[26] C.-F. Yung, Y.-F. Li, and H.-T. Sheu, "H_∞ filtering and solution bound for non-linear systems," *Int. J. Control*, vol. 74, no. 6, pp. 565–570, 2001.

[27] W. Chen and L. C. Jiao, "Finite-time stability theorem of stochastic nonlinear systems," *Automatica*, vol. 46, no. 12, pp. 2105–2108, 2010.

[28] W. Zhang and B.-S. Chen, "State feedback H_∞ control for a class of nonlinear stochastic systems," *SIAM J. Control Optim.*, vol. 44, no. 6, pp. 1973–1991, 2006.

[29] S. Xu and T. Chen, "H_∞ output feedback control for uncertain stochastic systems with time-varying delays," *Automatica*, vol. 40, no. 12, pp. 2091–2098, 2004.

[30] H. G. Tanner, A. Jadbabaie, and G. J. Pappas, "Stable flocking of mobile agents, part I: fixed topology," in *Proc. 2003 IEEE Conf. Decision Control.*, 2003, pp. 2010–2015.

[31] C. W. Wu, *Synchronization in Complex Networks of Nonlinear Dynamical Systems*, vol. 76. Singapore: World Scientific, 2007.

[32] W. Lu and T. Chen, "New approach to synchronization analysis of linearly coupled ordinary differential systems," *Physica D, Nonlinear Phenomena*, vol. 213, no. 2, pp. 214–230, 2006.

[33] W. Guo, F. Austin, S. Chen, and W. Sun, "Pinning synchronization of the complex networks with non-delayed and delayed coupling," *Phys. Lett. A*, vol. 373, no. 17, pp. 1565–1572, 2009.

[34] S. P. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, vol. 15. Philadelphia, PA, USA: SIAM, 1994.

[35] W. Yu, G. Chen, M. Cao, and J. Kurths, "Second-order consensus for multiagent systems with directed topologies and nonlinear dynamics," *IEEE Trans. Syst., Man, Cybern., B, Cybern.*, vol. 40, no. 3, pp. 881–891, Jun. 2010.

[36] K. Mathiyalagan, R. Sakthivel, and S. M. Anthoni, "Exponential stability result for discrete-time stochastic fuzzy uncertain neural networks," *Phys. Lett. A*, vol. 376, no. 8, pp. 901–912, 2012.

[37] S. S. Ge and C. Wang, "Direct adaptive NN control of a class of nonlinear systems," *IEEE Trans. Neural Netw.*, vol. 13, no. 1, pp. 214–221, Jan. 2002.

Guoxing Wen received the M.S. degree in applied mathematics from the 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 September 2015 to September 2016. He is currently a Lecturer with the College of Science, Binzhou University, Binzhou, China. His current research interests include adaptive neural network control, optimal control, and multiagent 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, both the Ph.D. degree from Purdue University, West Lafayette, IN, USA, both in electrical engineering, in 1985 and 1988, respectively.

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.

Yan-Jun Liu received the B.S. degree in applied mathematics and the M.S. degree in control theory and control engineering from the Shenyang University of Technology, Shenyang, China, in 2001 and 2004, respectively, and the Ph.D. degree in control theory and control engineering from the Dalian University of Technology, Dalian, China, in 2007.

He is currently a Professor with the College of Science, Liaoning University of Technology, Jinzhou, China. His current research interests include adaptive fuzzy control, nonlinear control, neural network control, and reinforcement learning.

报告编号：BZU20180033

文献检索证明

作者姓名：文国兴

作者单位：滨州学院

该作者 2018 年发表在期刊《IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS》第 65 卷第 7 期的论文《Formation Control With Obstacle Avoidance for a Class of Stochastic Multiagent Systems》被 SCIE（SCI-EXPANDED）收录。该期刊 2017 年 JCR 的影响因子是 7.05，JCR 分区见附件。在中科院 2016 年 JCR 分区表中位于大类分区 1 区。

检索结果见附件。

证明单位：滨州学院图书馆

证明人：郭春霞

2018 年 10 月 26 日

第 1 条，共 1 条
标题: Formation Control With Obstacle Avoidance for a Class of Stochastic Multiagent Systems
作者: Wen, GX (Wen, Guoxing); Chen, CLP (Chen, C. L. Philip); Liu, YJ (Liu, Yan-Jun)
来源出版物: IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS 卷: 65 期: 7 页: 5847-5855 DOI: 10.1109/TIE.2017.2782229 出版年: JUL 2018
Web of Science 核心合集中的 "被引频次": 2
被引频次合计: 2
入藏号: WOS:000427132300061
文献类型: Article
地址: [Wen, Guoxing] Binzhou Univ, Coll Sci, Binzhou 256600, Peoples R China.
[Chen, C. L. Philip] Univ Macau, Fac Sci & Technol, Dept Comp & Informat Sci, Macau 99999, Peoples R China.
[Chen, C. L. Philip] DalianMaritime Univ, Dalian 116026, Peoples R China.
[Chen, C. L. Philip] Chinese Acad Sci, Inst Automat, State Key Lab Management & Control Complex Syst, Beijing 100080, Peoples R China.
[Liu, Yan-Jun] Liaoning Univ Technol, Coll Sci, Jinzhou 121001, Peoples R China.
通讯作者地址: Wen, GX (通讯作者)， Binzhou Univ, Coll Sci, Binzhou 256600, Peoples R China.
电子邮件地址: gxwen@live.cn; philip.chen@ieee.org; liuyanjun@live.com
ISSN: 0278-0046
eISSN: 1557-9948
来源出版物页码计数: 9

关闭 打印
Web of Science
第 1 页 (记录 1 -- 1)
◀[1]▶

IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS

impact factor

7.05 7.623
2017 5 年

JCR® 类别	类别中的排序	JCR 分区
AUTOMATION & CONTROL SYSTEMS	2/61	Q1
ENGINEERING, ELECTRICAL & ELECTRONIC	13/260	Q1
INSTRUMENTS & INSTRUMENTATION	1/61	Q1

数据来自第 2017 版 Journal Citation Reports

出版商
IEEE-INST ELECTRICAL ELECTRONICS ENGINEERS INC, 445 HOES LANE, PISCATAWAY, NJ 08855-4141 USA

ISSN: 0278-0046
eISSN: 1557-9948

研究领域
Automation & Control Systems
Engineering
Instruments & Instrumentation

关闭窗口

Neural Network-Based Adaptive Leader-Following Consensus Control for a Class of Nonlinear Multiagent State-Delay Systems

Guoxing Wen, C. L. Philip Chen, *Fellow, IEEE*, Yan-Jun Liu, and Zhi Liu

***Abstract*—Compared with the existing neural network (NN) or fuzzy logic system (FLS) based adaptive consensus methods, the proposed approach can greatly alleviate the computation burden because it needs only to update a few adaptive parameters online. In the multiagent agreement control, the system uncertainties derive from the unknown nonlinear dynamics are counteracted by employing the adaptive NNs; the state delays are compensated by designing a Lyapunov–Krasovskii functional. Finally, based on Lyapunov stability theory, it is demonstrated that the proposed consensus scheme can steer a multiagent system synchronizing to the predefined reference signals. Two simulation examples, a numerical multiagent system and a practical multimanipulator system, are carried out to further verify and testify the effectiveness of the proposed agreement approach.**

***Index Terms*—Consensus control, external disturbance, neural networks (NNs), nonlinear multiagent systems, state delay.**

I. Introduction

IN RECENT decades, multiagent system control has become an attractive and active research topic because of its wide applications in various fields, such as flocking, distributed sensor networks, unmanned aerial vehicle formation, etc. [1]–[5]. In multiagent control community, the consensus is one of the most fundamental research topics [6]. Roughly speaking, consensus control of a multi-agent system is all agents to be synchronized to a common state by a control protocol based on the neighbor agents' information. Usually, consensus control can be divided into two classes that are leaderless consensus and leader-following consensus, of which leader-following consensus is that all agents are synchronized by following a common reference signal. In the recent decades, many excellent consensus methods have been reported (see [6]–[8]). In [6], the consensus problem is analyzed for the multiagent dynamics with fixed and switching topologies, and two classes of consensus methods are introduced for the communication networks without and with time-delays. In [7], the leader-following consensus control is addressed for the higher order multiagent systems, and every agent's controller is constructed using the local information. The agreement control design is achieved by integrating the algebraic graph theory, Riccati inequality and Lyapunov stability analysis together. In [8], two novel distributed adaptive dynamic consensus protocols are proposed, in which one protocol assigns an adaptive coupling weight to each edge in communication graph and the other uses an adaptive coupling weight for each node. Although these consensus methods mentioned in [6]–[8] have good control performance for the linear systems, they are difficult to be generalized to the nonlinear systems.

It is well known that the nonlinear model can describe system dynamic really. In recent years, a few nonlinear consensus controls have been reported and received considerable attentions, for instance, [9]–[12], but these methods do not consider any time delay and external disturbance. In fact, most existing consensus methods about the time delay problem are only focused on the linear multiagent systems [13]–[15]. For the nonlinear multiagent systems, it is still an unexplored research topic. In practical engineering systems, the state delays and external disturbances are often encountered, even they can degrade the system performance and possibly cause the system instability, especially when the time delays and external disturbances are not exactly known. Two main methods dealt with the state delay problems, Lyapunov–Krasovskii functional and Lyapunov–Razumikhin function methods, have been well-developed for the tracking or regulation control of nonlinear systems, where Lyapunov–Krasovskii functional method is a simple and convenient means [16]–[18]. Because most real multiagent systems contain inherent state delays, it is very necessary to consider the time delay problem for the multiagent system control design.

Neural networks (NNs) or fuzzy logical systems (FLSs) have become an effective and powerful tools in the nonlinear system modeling and controlling owing to the excellent approximation and learning abilities. In the past decades, a

Manuscript received March 5, 2016; revised June 20, 2016 and August 23, 2016; accepted August 29, 2016. Date of publication October 11, 2016; date of current version July 17, 2017. This work was supported in part by the Doctoral Scientific Research Staring Fund of Binzhou University, under Grant 2016Y14, and in part by the National Natural Science Foundation of China under Grant 61622303. This paper was recommended by Associate Editor L. Zhang.

G. Wen is with the Department of Mathematics, 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 256600, China, and also with UMacau Research Institute, Zhuhai 519080, China (e-mail: Philip.Chen@ieee.org).

Y.-J. Liu is with the College of Science, Liaoning University of Technology, Jinzhou 121001, China (e-mail: liuyanjun@live.com).

Z. Liu is with the Department of Automation, Guangdong University of Technology, Guangzhou 510006, China (e-mail: lz@gdut.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCYB.2016.2608499

great number of NN or FLS-based nonlinear control methods have been published [19]–[26], in which [19]–[22] are based on NN and [23]–[26] are based on FLS. In [19] and [22], the NN dynamic surface technique-based adaptive tracking control is developed for the strict feedback systems. In [20], the NN reinforcement learning is integrated into the output control of nonlinear strict feedback systems. In [23], the FLS observer is constructed for the nonlinear systems to estimate the unmeasured states. It is worth mentioning that a few NN or FLS-based adaptive consensus approaches are proposed, and they have received increasing attention [27]–[29] recently. However, these adaptive consensus methods, for obtaining the desired approximation accuracy, often require a large number of adaptive parameters. Therefore, if these consensus methods are applied to the practical engineering systems, it will result in heavy online computational burden and be implemented difficultly.

Motivated by above discussion, this paper addresses leader-following consensus control for nonlinear multiagent time delay systems. The main contributions in this paper are listed as follows.

1) Compared with the existing research results, the proposed consensus control method can alleviate the computation burden because only a small number of adaptive parameters are updated. It means that the proposed agreement method can reduce the running cost and easily apply to the practical multiagent engineering.
2) By employing the Lyapunov–Krasovskii functionals, the impact coming from the unknown state delays is compensated.
3) By employing the adaptive NN-based approximation techniques, the difficulties coming from the unknown nonlinear dynamics and external disturbances are well overcome.

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

1) R represents real number; R^n denotes real m-dimensional vector space; $R^{n\times m}$ is $n\times m$-dimensional matrix space; Ω is a subset of R^n; and I_n is $n\times n$ identity matrix.
2) $\|\cdot\|$ represents 2-norm of vector; $\|\cdot\|_F$ represents Frobenius norm of matrix.
3) If there is no special explanation, T represents the transposition symbol.

II. Problem Formulation and Preliminaries

A. System Description and Assumptions

Consider a class of nonlinear multiagent systems modeled by the following differential equation:

$$\dot{x}_i(t) = f_i(x_i(t)) + g_i(x_i(t))u_i(t) + h_i(x_i(t-\tau_i)) + d_i(t, x_i)$$
$$i = 1, 2, \ldots, n \tag{1}$$

where $x_i(t) \in R^m$ is the state vector; $u_i(t) \in R^m$ is the control input; $f_i(\cdot)$, $h_i(\cdot) : R^m \to R^m$ are the smooth vector-value nonlinear functions with the uncertainties. The control gain matrix $g_i(\cdot) : R^m \to R^{m\times m}$ is a strictly either positive or negative definite, of which each element is an unknown nonlinear function or constant; τ_i is the unknown time delay; $d_i(t, x_i) \in R^m$ is the unknown external disturbance.

The desired reference is described by the following equation, it is taken as the leader of nonlinear multiagent system (1) in the control design:

$$\dot{x}_l(t) = f_l(t) \tag{2}$$

where $x_l \in R^m$ is the leader's state and $f_l(t) \in R^m$ is a smooth vector-value function.

In this paper, the control task is designing a consensus control scheme for the multiagent system (1) such that all agents can synchronously track the desired trajectory to a desired accuracy, i.e., $\lim_{t\to\infty} \|x_i(t) - x_l(t)\| = 0$ and $i = 1, 2, \ldots, n$.

Assumption 1: For the unknown time delays τ_i, $i = 1, 2, \ldots, n$, there exists a positive known constant $\tau_{\max}$ such that $\tau_i \le \tau_{\max}$, $i = 1, \ldots, n$ [17].

Assumption 2: The nonlinear function $f_l(t) \in R^m$ of the leader's dynamic is bounded, i.e., $\|f_l(t)\| < \alpha$, where α is a positive constant [17].

In Assumption 2, the constant α is only for stability analysis, its actual value is unnecessary to be known.

Assumption 3: For these terms $h_i(x_i(t))$, $i = 1, 2, \ldots, n$, there exist the known smooth functions $\varrho_i(x_i(t))$ to satisfy $\|h_i(x_i(t))\| \le \varrho_i(x_i(t))$ [30].

Assumption 4: For the disturbance dynamics $d_i(t, x_i)$, $i = 1, 2, \ldots, n$, there exists the unknown continuous functions $p_i(x_i(t))$ satisfying $\|d_i(t, x_i)\| \le p_i(x_i(t))$ [17].

Assumption 5: There exist two positive or negative constants, $\underline{g}_i$ and $\bar{g}_i$, such that $\underline{g}_i \le \lambda_1(g_i(x_i)), \ldots, \lambda_m(g_i(x_i)) \le \bar{g}_i$, $i = 1, \ldots, n$, where $\lambda_1(g_i(x_i)), \ldots, \lambda_m(g_i(x_i))$ are all eigenvalues of the matrix function $g_i(x_i(t))$. Without loss of generality, it is assumed that $\lambda_1(g_i(x_i)), \ldots, \lambda_m(g_i(x_i)) \ge \underline{g}_i > 0$, $i = 1, \ldots, n$ [17].

B. Algebraic Graph Theory

Let $G := (\Upsilon, \varepsilon, A)$ denote a weight digraph, which is used to describe the communication topology of the multiagent system (1), where $\Upsilon := \{\upsilon_1, \upsilon_2, \ldots, \upsilon_n\}$ denotes the node set; $\varepsilon \subseteq \Upsilon \times \Upsilon$ denotes the edge set; and $A = [a_{ij}]$ denotes the adjacency matrix. The node υ_i represents the ith agent. Let $\epsilon_{ij} = (\upsilon_i, \upsilon_j)$ denote an edge of the weight graph G, $\epsilon_{ij} = (\upsilon_i, \upsilon_j) \in \varepsilon$ if and only if there is a communication from agent j to agent i. We say node υ_j is a neighbor of node υ_i if the edge $\epsilon_{ij} = (\upsilon_i, \upsilon_j) \in \varepsilon$. The neighbor set of node υ_i is described by $N_i := \{\upsilon_j \| (\upsilon_i, \upsilon_j) \in \varepsilon\}$. The adjacency element a_{ij} corresponding to the edge ϵ_{ij} denotes the communication quality between the agents i and j, i.e., $\epsilon_{ij} \in \varepsilon \Leftrightarrow a_{ij} > 0$; otherwise $a_{ij} = 0$. A weight graph G is called undirected if and only if $a_{ij} = a_{ji}$. An undirected graph implies that node υ_j is a neighbor of node υ_i if and only if node υ_i is also a neighbor of node υ_j. Laplacian matrix $L = [l_{ij}] \subset R^{n\times n}$ for the weight graph G is defined as

$$L := C - A \tag{3}$$

where $C = \mathrm{diag}\{c_1, c_2, \ldots, c_n\}$, $c_i = \sum_{j=1}^{n} a_{ij}$. Obviously, $\mathbf{1}_n = [1, 1, \ldots, 1]^T \in R^n$ is an eigenvector associated with the eigenvalue $\lambda = 0$.

Let $B := \text{diag}\{b_1, b_2, \ldots, b_n\}$ denote the communication weight matrix between agents and leader. $b_i > 0$ if and only if there exists the information exchange between agent i and leader, otherwise $b_i = 0$. It is stipulated that at least one agent connects with leader, i.e., $b_1 + b_2 + \cdots + b_n > 0$.

A sequence of edges of a graph G is called a path if it is the form that $(\upsilon_i, \upsilon_{i_1}), (\upsilon_{i_1}, \upsilon_{i_2}), \ldots, (\upsilon_{i_l}, \upsilon_j)$. An undirected graph is called connected if there is a path for any a pair of distinct nodes. For a connected graph, all nonzero eigenvalues of L are non-negative, and 0 is a simple eigenvalue of L [31].

C. Radial Basis Function Neural Networks and Function Approximation

It has been proven that the radial basis function NNs (RBFNNs) have the universal approximation and learning abilities. Any unknown smooth function $\psi(x) : R^n \to R^m$ can be approximated by RBFNNs in the following form:

$$\hat{\psi}(x) = W^T S(x)$$

where $x \in \Omega_x \subset R^n$, Ω_x is a compact set, $W \in R^{q\times m}$ is the adjustable weight matrix with the number of neurons q, $S(z) = [s_1(x), \ldots, s_q(x)]^T$ is the basis function vector, $s_i(x) = \exp(-(x - \nu_i)^T (x - \nu_i)/\varphi_i^2), i = 1, 2, \ldots, q$, $\nu_i = [\nu_{i1}, \nu_{i2}, \ldots, \nu_{in}]^T$ is the center of the receptive field, φ_i is the width of the Gaussian function.

It is well known that RBFNNs can approximate a continuous function to any desired accuracy by making the neuron number q large enough and choosing the design parameters appropriately. For the smooth function $\psi(x)$, there exists an ideal weights W^* such that

$$\psi(x) = W^{*T} S(x) + \varepsilon(x) \tag{4}$$

where $\varepsilon(x) \in R^m$ is the approximation error to satisfy $\|\varepsilon(x)\| \leq \delta$, δ is a positive constant. The NN approximation error indicates the minimum possible deviation between the ideal approximation $W^* S(x)$ and the smooth unknown function $\psi(x)$.

In fact, the ideal NN weight matrix W^* is an "artificial" quantity just for analysis purposes and it needs to be estimated in control design [32]. W^* is defined as

$$W^* := \arg \min_{W\in R^{p\times m}} \left\{ \sup_{x\in\Omega_z} \|\psi(x) - WS(x)\| \right\}. \tag{5}$$

D. Supporting Lemmas

Lemma 1 [31]: An undirected graph G is connected if and only if its Laplacian is irreducible.

Lemma 2 [33]: Let $D = [d_{ij}] \in R^{n\times n}$ be an irreducible matrix such that $d_{ij} = d_{ji} \leq 0$ for $i \neq j$ and $d_{ii} = -\sum_{j=1}^{n} l_{ij}$ for $i = 1, 2, \ldots, n$. Then all eigenvalues of the matrix $\tilde{D} = \begin{bmatrix} d_{11}+\theta_1 & \cdots & d_{1n} \\ \vdots & \ddots & \vdots \\ d_{n1} & \cdots & d_{nn}+\theta_n \end{bmatrix}$ are positive, where $\theta_1, \theta_2, \ldots, \theta_n$ are non-negative constants and $\theta_1 + \theta_2 + \cdots + \theta_n > 0$.

Lemma 3 [34]: Let $R(t) \in R$ be a continuous positive function with bounded initial value $R(0)$. If the inequality holds that $\dot{R}(t) \leq -\beta R(t) + \gamma$, where β and γ are positive constants, then the following inequality is held:

$$R(t) \leq R(0)\mathrm{e}^{-\beta t} + \frac{\gamma}{\beta}\left(1 - \mathrm{e}^{-\beta t}\right). \tag{6}$$

III. MAIN RESULTS

In the work, the interconnection graph G of the nonlinear multiagent system (1) is assumed to be an undirected connected graph. Define the tracking error variable between the ith agent and leader as $\bar{\zeta}_i(t) = x_i(t) - x_l(t)$. Based on the system dynamic equations (1) and (2), the error dynamics are obtained as

$$\begin{aligned} \dot{\bar{\zeta}}_i(t) &= f_i(x_i(t)) + g_i(x_i(t))u_i(t) + h_i(x_i(t-\tau_i)) \\ &\quad + d_i(t, x_i) - f_l(t) \\ i &= 1, 2, \ldots, n. \end{aligned} \tag{7}$$

Define the consensus error vector for the ith agent as

$$\begin{aligned} e_i(t) &= \sum_{j\in N_i} \left(a_{ij}\left(x_i(t) - x_j(t)\right) + b_i(x_i(t) - x_l(t)) \right) \in R^m \\ i &= 1, 2, \ldots, n \end{aligned} \tag{8}$$

where a_{ij} is the ith row and jth column element of the adjacency matrix A; $b_i \geq 0$ is connection weight between the ith agent and leader. Using the tracking error variable $\bar{\zeta}(t)_i = x_i(t) - x_l(t)$, the consensus error vector (8) can be rewritten as

$$e_i(t) = \sum_{j\in N_i} \left(a_{ij}\left(\bar{\zeta}_i(t) - \bar{\zeta}_j(t)\right) + b_i \bar{\zeta}_i(t) \right) \in R^m. \tag{9}$$

Define a scalar function as

$$V_1(t) = \frac{1}{2}\bar{\zeta}^T(t)\left(\tilde{L} \otimes I_m\right)\bar{\zeta}(t) \tag{10}$$

where $\bar{\zeta} = [\bar{\zeta}_1^T(t), \bar{\zeta}_2^T(t), \ldots, \bar{\zeta}_n^T(t)]^T \in R^{nm}$; $\tilde{L} = L + B$, $B = \text{diag}\{b_1, b_2, \ldots, b_n\}$. According to Lemma 2, $V_1(t)$ is a positive definite function.

Because $\tilde{L}$ is a symmetrical positive definite matrix, it has n positive eigenvalues that is denoted by $\lambda_1, \lambda_2, \ldots, \lambda_n$. Let $\chi_{11}, \ldots, \chi_{1m}, \chi_{21}, \ldots, \chi_{2m}, \ldots, \chi_{n1}, \ldots, \chi_{nm}$ denote the eigenvectors of the positive definite matrix $\tilde{L} \otimes I_m$ corresponding to the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ respectively, they can be chosen as a set of orthogonal bases of R^{nm}. Let $M = [\chi_{11}, \ldots, \chi_{nm}] \in R^{nm\times nm}$, then the equation that $M^T M = M M^T = I_{nm}$ is held.

Based on above analysis, the scalar function, $V_1(t)$, can be rewritten as

$$\begin{aligned} V_1(t) &= \frac{1}{2}\bar{\zeta}^T(t)\left(\tilde{L} \otimes I_m\right)\bar{\zeta}(t) = \frac{1}{2}\bar{\zeta}^T(t) M^T \Lambda M \bar{\zeta}(t) \\ &= \frac{1}{2}\bar{\zeta}^T(t) M^T \Lambda \Lambda^{-1} \Lambda M \bar{\zeta}(t) \\ &= \frac{1}{2}\bar{\zeta}^T(t) M^T \Lambda M M^T \Lambda^{-1} M M^T \Lambda M \bar{\zeta}(t) \\ &= \frac{1}{2}\bar{\zeta}^T(t)\left(\tilde{L} \otimes I_m\right)^T M^T \Lambda^{-1} M \left(\tilde{L} \otimes I_m\right)\bar{\zeta}(t) \\ &= \frac{1}{2} e^T(t) \Delta e(t) \end{aligned} \tag{11}$$

where $\otimes$ is Kronecker product, $e(t) = [e_1^T(t), \dots e_n^T(t)]^T \in R^{nm}$, $\Lambda = \mathrm{diag}\{\lambda_1 I_m, \lambda_2 I_m, \dots, \lambda_n I_m\}$ and $\Delta = M^T \Lambda^{-1} M$.

Based on (11), the following one can be obtained:

$$\frac{\lambda_{\min}(\Delta)}{2}\|e(t)\|^2 \le V_1(t) \le \frac{\lambda_{\max}(\Delta)}{2}\|e(t)\|^2 \tag{12}$$

where $\lambda_{\min}(\Delta)$ and $\lambda_{\max}(\Delta)$ are the smallest and largest eigenvalues of matrix Δ, respectively.

Taking time derivative of $V_1(t)$ along (7) is

$$\begin{aligned}\dot{V}_1(t) &= \bar{\zeta}^T(t)\big(\tilde{L}\otimes I_m\big)\dot{\bar{\zeta}}(t) = \sum_{i=1}^n e_i^T(t)\dot{\bar{\zeta}}_i(t)\\ &= \sum_{i=1}^n \big(e_i^T(t)(g_i(x_i(t))u_i(t) + f_i(x_i(t)) + h_i(x_i(t-\tau_i))\\ &\quad + d_i(t,x_i) - f_l(t))\big).\end{aligned} \tag{13}$$

Based on Assumptions 2–4, the following results can be obtained by applying Cauchy inequality that $(\sum_{i=1}^n x_i y_i)^2 \le (\sum_{i=1}^n x_i^2)(\sum_{i=1}^n y_i^2)$:

$$-e_i^T(t)f_l(t) \le \|e_i(t)\|\|f_l(t)\| \le \alpha\|e_i(t)\| \tag{14}$$

$$e_i^T(t)h_i(x_i(t-\tau_i)) \le \|e_i(t)\|\varrho_i(x_i(t-\tau_i)) \tag{15}$$

$$e_i^T(t)d_i(t,x) \le \|e_i(t)\|p_i(x(t)). \tag{16}$$

Substituting (14)–(16) into (13), the following inequality can be yielded:

$$\begin{aligned}\dot{V}_1(t) \le \sum_{i=1}^n &\big(e_i^T(t)g_i(x_i(t))u_i(t) + e_i^T(t)f_i(x_i(t)) + \alpha\|e_i(t)\|\\ &+ \|e_i(t)\|\varrho_i(x_i(t-\tau_i)) + \|e_i(t)\|p_i(x(t))\big).\end{aligned} \tag{17}$$

Remark 1: In inequality (16), the unknown time delay τ_i is an obstacle for the controller design. Although the scalar function $\varrho_i(x_i(t))$ is known, $\varrho_i(x_i(t-\tau_i))$ will become undetermined owing to the unknown delay τ_i. Because the unknown time delay function $\varrho_i(x_i(t-\tau_i))$ and the consensus error $e_i(t)$ are merged together, the problem become more complex for the control design. Therefore, these related terms need to be transformed to the form that the uncertain time delay term $\varrho_i(x_i(t-\tau_i))$ and the consensus error $e_i(t)$ are separated.

Applying Young's inequality, $ab \le (a^2/2) + (b^2/2)$, the following results can be yielded:

$$\alpha\|e_i(t)\| \le \frac{\|e_i(t)\|^2}{2} + \frac{\alpha^2}{2} \tag{18}$$

$$\|e_i(t)\|\varrho_i(x_i(t-\tau_i)) \le \frac{\|e_i(t)\|^2}{2} + \frac{\varrho_i^2(x_i(t-\tau_i))}{2} \tag{19}$$

$$\|e_i(t)\|p_i(x(t)) \le \frac{\beta_i^2}{2} + \frac{\|e_i(t)\|^2 p_i^2(x_i(t))}{2\beta_i^2} \tag{20}$$

where β_i is a positive constant. Using (18)–(20), the inequality (17) can be rewritten as

$$\begin{aligned}\dot{V}_1(t) \le \sum_{i=1}^n &\Bigg(e_i^T(t)g_i(x_i(t))u_i(t) + \|e_i(t)\|^2 + e_i^T(t)f_i(x_i(t))\\ &+ \frac{1}{2\beta_i^2}\|e_i(t)\|^2 q_i^2(x_i(t)) + \frac{1}{2}\varrho_i^2(x_i(t-\tau_i))\Bigg)\\ &+ \frac{1}{2}\left(\sum_{i=1}^n \beta_i^2 + n\alpha^2\right).\end{aligned} \tag{21}$$

In inequality (21), $e_i(t)$ and $\varrho_i(x_i(t-\tau_i))$ are separated. Then the following Lyapunov–Krasovskii functional is used to eliminate the difficulties in the control design come from the unknown time delay $\tau_i, i = 1, \dots, n$:

$$V_2(t) = \frac{1}{2}\sum_{i=1}^n \int_{t-\tau_i}^t \varrho_i^2(x_i(s))ds. \tag{22}$$

Taking time derivative of $V_2(t)$ is

$$\dot{V}_2(t) = \frac{1}{2}\sum_{i=1}^n \varrho_i^2(x_i(t)) - \frac{1}{2}\sum_{i=1}^n \varrho_i^2(x_i(t-\tau_i)). \tag{23}$$

Obviously, $\dot{V}_2(t)$ can compensate the uncertainties of the inequality (21) derived from the time-delay τ_i, and thus the design difficulty is eliminated. Choose Lyapunov function candidate for the dynamic systems (1) as $V_e(t) = V_1(t) + V_2(t)$, based on (21) and (23), its time derivative is

$$\begin{aligned}\dot{V}_e(t) &= \dot{V}_1(t) + \dot{V}_2(t)\\ &\le \sum_{i=1}^n \Bigg(e_i^T(t)g_i(x_i)u_i + \|e_i(t)\|^2 + e_i^T(t)f_i(x_i)\\ &\qquad + \frac{1}{2\beta_i^2}\|e_i(t)\|^2 q_i^2(x_i) + \frac{1}{2}\varrho_i^2(x_i)\Bigg)\\ &\quad + \frac{1}{2}\left(\sum_{i=1}^n \beta_i^2 + n\alpha^2\right)\\ &= \sum_{i=1}^n \left(e_i^T(t)g_i(x_i)u_i + \|e_i(t)\|^2 + e_i^T(t)Q_i(z_i) + \frac{1}{2}\varrho_i^2(x_i)\right)\\ &\quad + \frac{1}{2}\left(\sum_{i=1}^n \beta_i^2 + n\alpha^2\right)\end{aligned} \tag{24}$$

where $Q_i(z_i) = f_i(x_i) + (1/2\beta_i^2)e_i(t)q_i^2(x_i)$; $z_i = \{x_i(t), e_i(t)\} \in \Omega_{z_i}$, Ω_{z_i} is a compact set.

Under the condition that all system functions are known, $V_e(t)$ can be chosen as Lyapunov function. In order to finish the control task, the desired controller $u_i, i = 1, \dots, n$ is constructed in the following:

$$u_i(t) = \begin{cases} -\mathrm{k}_i(t)e_i(t) - g_i^{-1}(x_i)Q_i(z_i) & \\ -\frac{1}{2\underline{g}_i}e_i^{-1}(t)\varrho_i^2(x_i) & e_i(t) \in \Omega_{\phi_i}^0\\ 0 & e_i(t) \in \Omega_{\phi_i}\end{cases} \tag{25}$$

where $k_i(t) \in R^+, i = 1, \dots, n$, $g_i^{-1}(x_i)$ is the inverse of matrix $g_i(x_i)$, $e_i^{-1}(t) = e_i(t)/\|e_i(t)\|^2$, $\Omega_{\phi_i} = \{z_i | \|e_i(t)\| < \phi_i\} \subset \Omega_{z_i}$, $\Omega_{\phi_i}^0 = \Omega_{z_i} - \Omega_{\phi_i}$, ϕ_i is an arbitrary small constant. $\Omega_{\phi_i}^0$ is also a compact set [17].

Remark 2: Because the term $(1/2)e_i^{-1}(t)\varrho_i^2(x_i)$ is not well defined at $e_i(t) = [0]_m$, the controller singularity problem may occur when the term $(1/2)e_i^{-1}(t)\varrho_i^2(x_i)$ is utilized to construct the consensus controller. Therefore, the boundedness of the control must be guaranteed. It need to be noted that the control objective has been achieved when $e_i(t) = [0]_m$, so relaxing the consensus tracking error converges to a "ball" is more practical than origin [35].

When $e_i \in \Omega^0_{\phi_i}$, substituting the controller (25) into (24), the following inequality is obtained:

$$\dot{V}_e \le \sum_{i=1}^{n}\Bigg(- k_i(t)e_i^T(t)g_i(x_i)e_i(t) + \|e_i(t)\|^2 - \frac{e_i^T(t)g_i(x_i)e_i(t)}{2\underline{g}_i\|e_i(t)\|^2}\varrho_i^2(x_i(t)) + \frac{1}{2}\varrho_i^2(x_i)\Bigg) + \frac{1}{2}\Bigg(\sum_{i=1}^{n}\beta_i^2 + n\alpha^2\Bigg). \quad (26)$$

Applying Assumption 5, above inequality can become

$$\dot{V}_e \le -\sum_{i=1}^{n}\Big(\underline{g}_i k_i(t) - 1\Big)\|e_i(t)\|^2 + \frac{1}{2}\Bigg(\sum_{i=1}^{n}\beta_i^2 + n\alpha^2\Bigg). \quad (27)$$

Let

$$k_i(t) = \frac{\gamma_i}{\underline{g}_i}\Bigg(\frac{\lambda_{\max}(\Delta)}{2} + \frac{1}{2\|e_i(t)\|^2}\int_{t-\tau_{\max}}^{t}\varrho_i^2(x_i(s))ds + \frac{1}{\gamma_i}\Bigg) \quad (28)$$

where $\gamma_i > 0$ is a design constant.

Substituting the controller gain (28) into (27) yields

$$\begin{aligned}\dot{V}_e &\le -\sum_{i=1}^{n}\frac{\gamma_i}{2}\lambda_{\max}(\Delta)\|e_i(t)\|^2 - \sum_{i=1}^{n}\frac{\gamma_i}{2}\int_{t-\tau_{\max}}^{t}\varrho_i^2(x_i(s))ds + \frac{1}{2}\Bigg(\sum_{i=1}^{n}\beta_i^2 + n\alpha^2\Bigg)\\ &\le -\frac{\gamma}{2}\lambda_{\max}(\Delta)\sum_{i=1}^{n}\|e_i(t)\|^2 - \frac{\gamma}{2}\sum_{i=1}^{n}\int_{t-\tau_{\max}}^{t}\varrho_i^2(x_i(s))ds + \eta\end{aligned} \quad (29)$$

where $\eta = (1/2)(\sum_{i=1}^{n}\beta_i^2 + n\alpha^2)$, $\gamma = \min\{\gamma_1, \gamma_2, \ldots, \gamma_n\}$.

Based on the inequality (12), the inequality (29) can be rewritten as

$$\dot{V}_e \le -\gamma V_1 - \frac{\gamma}{2}\sum_{i=1}^{n}\int_{t-\tau_{\max}}^{t}\varrho_i^2(x_i(s))ds + \eta \quad (30)$$

where $\tau_{\max}$, $\varrho_i(x_i(t))$, $i = 1, \ldots, n$ are known from Assumptions 1 and 3, so the term $(1/2)\sum_{i=1}^{n}\int_{t-\tau_{\max}}^{t}\varrho_i^2(x_i(s))ds$ does not contain any uncertainty. Because $\varrho_i^2(x_i(t)), i = 1, \ldots, n$ is positive, the following inequality holds:

$$\int_{t-\tau_{\max}}^{t}\varrho_i^2(x_i(s))ds \ge \int_{t-\tau_i}^{t}\varrho_i^2(x_i(s))ds. \quad (31)$$

The inequality (30) can be rewritten as

$$\dot{V}_e \le -\gamma V_1 - \gamma V_2 + \eta = -\gamma V_e + \eta. \quad (32)$$

Applying Lemma 3, the following result can be obtained:

$$V_e(t) \le \frac{\eta}{\gamma} + \bigg(V_e(0) - \frac{\eta}{\gamma}\bigg)\mathrm{e}^{-\gamma t}. \quad (33)$$

The inequality (33) implies the tracking error $\bar{\zeta}_i(t)$, $i = 1, \ldots, n$ can be decreased to small enough by choosing the appropriate design parameters γ_i, $i = 1, \ldots, n$.

Because $f_i(\cdot)$, $p_i(\cdot)$ are completely unknown so that $Q_i(z_i)$ are also unknown, the proposed controller (25) cannot be applied to the multiagent system (1). On the other hand, because $Q_i(z_i)$ is continuous and well-defined on the compact set $\Omega^0_{\phi_i}$, $Q_i(z_i)$ can be approximated to a desired accuracy by RBFNNs in the following form:

$$Q_i(z_i) = W_i^* S_i(z_i) + \varepsilon_i(z_i) \quad (34)$$

where $W_i^* \in R^{m\times q_i}$ is the ideal NN weight matrix, q_i denotes the number of neurons, $S_i(z_i) \in R^{q_i}$ are the basis functions, $\varepsilon_i \in R^m$ is approximation error satisfying $\|\varepsilon_i\| \le \delta_i$, and δ_i is a positive constant.

Based on above analysis, the adaptive consensus control laws for the nonlinear multiagent system (1) are designed as

$$u_i(t) = \begin{cases} -k_i(t)e_i(t) - \dfrac{\xi_i}{\underline{g}_i}\hat{w}_i(t)\|S_i(z_i)\|^2 e_i(t) & \\ -\dfrac{1}{2\underline{g}_i}e_i^{-1}(t)\varrho_i^2(x_i) & e_i(t) \in \Omega^0_{\phi_i} \\ 0 & e_i(t) \in \Omega_{\phi_i}\end{cases} \quad (35)$$

where $\hat{w}_i(t)$ is the estimation of the unknown adaptive constant w_i^*, $w_i^* = \|W_i^*\|_F^2$.

The adaption laws are designed as

$$\dot{\hat{w}}_i(t) = \kappa_i\Big(\xi_i\|S_i(z_i)\|^2\|e_i(t)\|^2 - \sigma_i\hat{w}_i(t)\Big) \quad (36)$$

where $\kappa_i, \xi_i, \sigma_i > 0, i = 1, \ldots, n$ are the design constants.

Remark 3: The $\sigma-$modification term $\sigma_i\hat{w}_i(t), i = 1, \ldots, n$ are used to improve robustness of the NN approximators, it can reduce a high gain control scheme for the case that estimates $\hat{w}_i(t)$ might shift to very high values [30].

Remark 4: In order to enhance the approximation accuracy, most existing approximator-based adaptive consensus control approaches require the neuron or fuzzy rule number large enough [27], [34] so that the online computational burden becomes very heavy. In the proposed controller, only a scalar adaptive parameter, a norm form of the NN weight matrix, is updated online for every agent. Because the computation burden is greatly reduced, it can be conveniently applied to the practical multiagent systems.

Remark 5: According to the updating law (36), for any bounded initial condition $\hat{w}_i(0) \ge 0$, if $\hat{w}_i(t) \le \xi_i\|S_i(z_i)\|^2\|e_i(t)\|^2/\sigma_i$, then $\dot{\hat{w}}_i(t) \ge 0$, thus $\hat{w}_i(t)$ is increased until $\hat{w}_i(t) = \xi_i\|S_i(z_i)\|^2\|e_i(t)\|^2/\sigma_i$; similarly, if $\hat{w}_i(t) > \xi_i\|S_i(z_i)\|^2\|e_i(t)\|^2/\sigma_i$, $\hat{w}_i(t)$ is decreased until $\hat{w}_i(t) = \xi_i\|S_i(z_i)\|^2\|e_i(t)\|^2/\sigma_i$. Therefore, (36) implies that $\hat{w}_i(t) \ge 0$ can be guaranteed for any bounded and positive initial condition of $\hat{w}_i(0)$.

The main result is summarized in the following theorem.

Theorem 1: Consider the nonlinear multiagent system (1) with the leader (2). If Assumptions 1–5 are satisfied, then the control protocol (35) with the adaptive NN updating law (36) can guarantee that the leader-following consensus is achieved for the bounded initial conditions $x_i(0)$ and $\hat{w}_i(0)$. The control gain, $k_i(t) = k_{i0} + k_{i1}(t)$, is designed as

$$k_{i0} \geq \frac{2}{\underline{g}_i}$$
$$k_{i1}(t) = \frac{\gamma_i}{2\underline{g}_i}\left(\lambda_{\max}(\Delta) + \frac{1}{\|e_i\|^2}\int_{t-\tau_{\max}}^{t} \varrho_i^2(x_i(s))ds\right) \quad (37)$$

where $\gamma_i > 0$ is a design constant.

Proof: see Appendix A ■

IV. Simulation Examples

In order to further demonstrate the effectiveness of the proposed consensus method, two simulation examples, a numerical multiagent example and a practical multimanipulator example, are carried out. In the two examples, every multiagent system has six agents. In the two agreement control design, the same RBF NN, the Laplacian matrix and time delays are chosen for simplicity, which are described in the following.

RBFNN is designed 36 nodes and centers ν_i evenly distribute in range $[-6, 6] \times [-6, 6]$, and the widths $\varphi_i = 2$. $S_i(z_i) = [s_1(z_i), \ldots, s_{36}(z_i)]^T$ with $s_j(z_i) = \exp[-(z_i - \nu_j)^T(z_i - \nu_j)/\varphi_j^2], j = 1, 2, \ldots, 36$, and the NN adaptive parameters $\hat{w}_i, i = 1, \ldots, 6$ are updated by (36) and the initial values $\hat{w}_i(0) = 0, i = 1, \ldots, 6$.

The communication weights between the six agents and leader are $B = \text{diag}\{0, 0.9, 0, 0.9, 0, 0\}$, and Laplacian matrix L is

$$L = \begin{bmatrix} 1.3 & -0.7 & 0 & 0 & 0 & -0.6 \\ -0.7 & 1.5 & -0.8 & 0 & 0 & 0 \\ 0 & -0.8 & 1.7 & -0.9 & 0 & 0 \\ 0 & 0 & -0.9 & 1.6 & -0.7 & 0 \\ 0 & 0 & 0 & -0.7 & 1.5 & -0.8 \\ -0.6 & 0 & 0 & 0 & -0.8 & 1.4 \end{bmatrix}.$$

The time delays are $\tau_1 = 1.4, \tau_2 = 1.5, \tau_3 = 1.6, \tau_4 = 1.7, \tau_5 = 1.8, \tau_6 = 1.9$, and $\tau_{\max} = 2$.

Example 1: The nonlinear multiagent time-delay systems are described as follows:

$$\frac{d}{dt}\begin{bmatrix} x_{i1} \\ x_{i2} \end{bmatrix} = \begin{bmatrix} x_{i2}(t)\sin(\alpha_{i1}x_{i1}(t)) \\ x_{i1}(t)\cos\left(\alpha_{i2}x_{i2}^2(t)\right) \end{bmatrix} + \left(1 + \cos(x_{i1}(t))\sin\left(x_{i2}^2(t)\right)\right)u_i + \begin{bmatrix} h_{i1}(x_i(t-\tau_i)) \\ h_{i2}(x_i(t-\tau_i)) \end{bmatrix} + \begin{bmatrix} d_{i1}(t, x_i) \\ d_{i2}(t, x_i) \end{bmatrix} \quad (38)$$

where $h_{i1}(x_i(t)) = \beta_{i1}x_{i1}(t)\cos(x_{i2}(t))$, $h_{i2}(x_i(t)) = \beta_{i2}x_{i2}(t)\sin(x_{i1}(t))$, $d_{i1}(t, x) = \gamma_{i1}x_{i1}^2(t)\cos(1.5t)$, and $d_{i2}(t, x) = \gamma_{i2}(x_{i1}^2(t) + x_{i2}^2(t))\sin(t)$. α_{i1}, α_{i2}, β_{i1}, β_{i2}, γ_{i1}, and γ_{i2} are shown in Tables I–III. The initial positions of six agents are $x_1(0) = (2, 0)^T$, $x_2(0) = (1.5, 1.5)^T$, $x_3(0) = (-1.5, 1)^T$, $x_4(0) = (-2, -1)^T$, $x_5(0) = (-1, -1.5)^T$, and $x_6(0) = (1, -2)^T$.

The leader's dynamic is described as

$$\frac{d(x_l(t))}{dt} = \begin{bmatrix} 8\sin(8t\pi) - 16\cos(8t\pi) \\ 4\cos(8t\pi) + 16\sin(8t\pi) \end{bmatrix}. \quad (39)$$

TABLE I
Values of $\alpha_{i1}, \alpha_{i2}, i = 1, \ldots, 6$

i	1	2	3	4	5	6
α_{i1}	0.7	-3.1	6.5	-11	9.5	9.5
α_{i2}	0.5	0.4	-5.5	-10.5	11.5	2

TABLE II
Values of $\beta_{i1}, \beta_{i2}, i = 1, \ldots, 6$

i	1	2	3	4	5	6
β_{i1}	1	3.2	-2.3	7	4.5	5.4
β_{i2}	1.3	2.5	0.8	-4.2	2.7	2.5

TABLE III
Values of $\gamma_{i1}, \gamma_{i2}, i = 1, \ldots, 6$

i	1	2	3	4	5	6
γ_{i1}	-2.1	3.4	2.8	6.1	3.7	7.8
γ_{i2}	1.6	0.8	6.3	5.2	-9	3.5

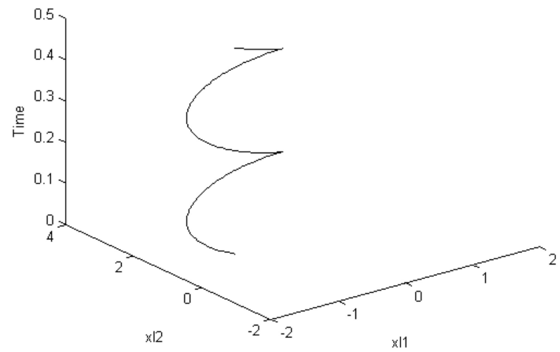


Fig. 1. Leader's trajectory.

Apparently, Assumption 3 can be satisfied by choosing $\varrho_i(x_i) = \sqrt{(\beta_{i1}x_{i1})^2 + (\beta_{i2}x_{i2})^2}$ and Assumption 4 can be satisfied by choosing $p_i(x_i) = \sqrt{\gamma_{i1}^2 x_{i1}^4 + \gamma_{i2}^2(x_{i1}^2 + x_{i2}^2)^2}$.

The adaptive NN controller u_i, $i = 1, \ldots, n$ are given by (35) and $\phi_i = 10^{-7}$, $i = 1, \ldots, 6$. The control gain $k_i(t) = k_{i0} + k_{i1}(t)$ is designed as $k_{i0} = 650$ and $k_{i1}(t)$ is given by (37) with $\gamma_i = 1/6$, $i = 1, \ldots, 6$. The correlation coefficients for the update laws are $\kappa_i = 0.4$, $\xi_i = 2$, $\sigma_i = 0.2$, $i = 1, \ldots, 6$.

Fig. 1 gives leader's trajectory. Figs. 2–4 show the simulation results applying the proposed consensus method to the systems (38). Figs. 3 and 4 display the leader-following agreement for the system (38) to be achieved.

Example 2: In this example, a multimanipulator example is carried out to test the effectiveness of the proposed consensus control scheme. The consensus control of the multimanipulator systems can be applied on many practical work occasions, for example, holding up a weight or loading a workpiece. The manipulator profile is shown in Fig. 5 and the system dynamic

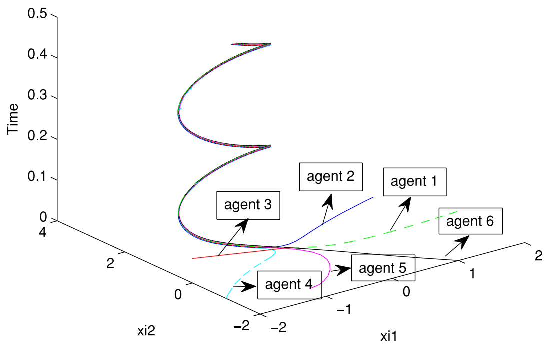


Fig. 2. Trajectory of six agents.

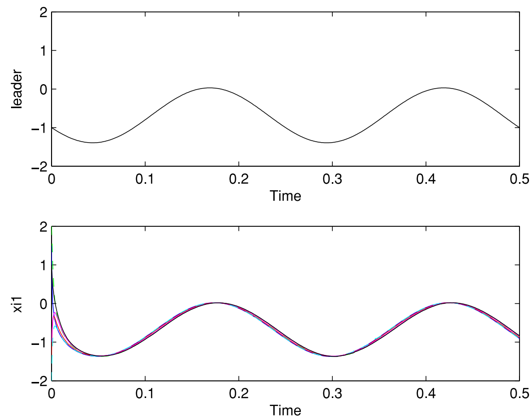


Fig. 3. First coordinate of leader and six agents.

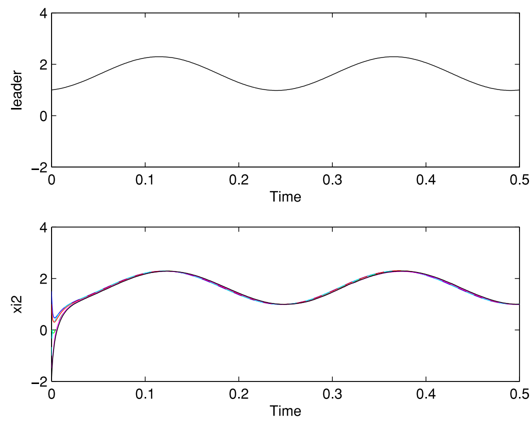


Fig. 4. Second coordinate of leader and six agents.

is described as

$$\dot{q}_{i1} = q_{i2}$$
$$M_i(q_{i1})\dot{q}_{i2} + V_i(q_{i1}, q_{i2})q_{i2} + G_i(q_{i1}) + f_i(q_{i2}(t - \tau_i)) = \zeta_i$$
$$i = 1, \ldots, 6 \quad (40)$$

where $q_{i1}, q_{i2} \in R^2$ denote the position and velocity state vectors of joints respectively; $M_i(q_{i1}) = \begin{bmatrix} M_{i11} & M_{i12} \\ M_{i21} & M_{i22} \end{bmatrix} \in R^{2\times 2}$ is the inertia matrix with $M_{i11} = (m_{i1} + m_{i2})r_{i1}^2 + 2m_{i2}r_{i1}r_{i2}\cos(q_{i12})$,

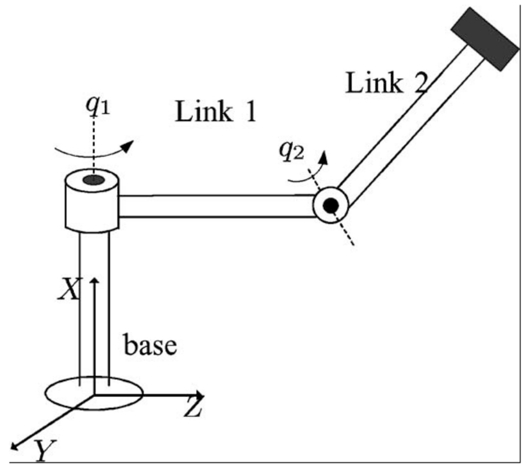


Fig. 5. Two-link revolute manipulator.

TABLE IV
VALUES OF $\alpha_{i1}, \alpha_{i2}, i = 1, \ldots, 6$

i	1	2	3	4	5	6
α_{i1}	11	6	-10	-9	12	13
α_{i2}	27	-14	22	18	-18	-16

TABLE V
VALUES OF $\beta_{i1}, \beta_{i2}, i = 1, \ldots, 6$

i	1	2	3	4	5	6
β_{i1}	4	2.6	-2.3	8	7	-5.4
β_{i2}	1.2	1.7	1.8	-4.2	-7	3.5

$M_{i12} = M_{i21} = m_{i2}r_{i2}^2 + m_{i2}r_{i1}r_{i2}\cos(q_{i12})$, $M_{i22} = m_{i2}r_{i2}^2$; $V_i(q_{i1}, q_{i2}) = \begin{bmatrix} V_{i11} & V_{i12} \\ V_{i21} & V_{i22} \end{bmatrix} \in R^{2\times 2}$ is the centripetal and Coriolis matrix with $V_{i11} = -m_{i2}r_{i1}r_{i2}\sin(q_{i12})q_{i22}$, $V_{i12} = -m_{i2}r_{i1}r_{i2}\sin(q_{i12})(q_{i21} + q_{i22})$, $V_{i21} = m_{i2}r_{i1}r_{i2}\sin(q_{i12})q_{i21}$, $V_{i22} = 0$; $G_i = (G_{i11}, G_{i12})^T \in R^2$ is gravitational vector with $G_{i1} = (m_{i1} + m_{i2})gr_{i1}\sin(q_{i11}) + m_{i2}gr_{i2}\sin(q_{i11} + q_{i12})$, $G_{i2} = m_{i2}gd_{i2}\sin(q_{i11} + q_{i12})$; $f_i(q_{i2}(t)) = (\alpha_{i1}q_{i21} + \beta_{i1}\text{sgn}(q_{i21}), \alpha_{i2}q_{i22} + \beta_{i2}\text{sgn}(q_{i22}))^T$ is the friction force vector, α_{i1}, α_{i2}, β_{i1}, β_{i2} are shown in Tables IV and V; $\zeta_i \in R^2$ is manipulator's torque input vector. The correlation parameters of the manipulators' dynamic are $g = 9.8\,\text{m/s}^2$, $r_{i1} = 1.6\,\text{m}$, $r_{i2} = 1.1\,\text{m}$, $m_{i1} = 1.1\,\text{kg}$ and $m_{i2} = 2.1\,\text{kg}$, $(i = 1, \ldots, 6)$. The initial joint velocities of six manipulators show in Table VI.

Leader dynamic can be described as

$$\dot{q}_{1d} = q_{2d}$$
$$\dot{q}_{2d} = [4\cos(4t) + 2\cos(6t), -2\sin(4t) - 4\cos(6t)]^T.$$

In the example, the consensus controllers are designed only for synchronizing the joint velocity q_{i2}, $i = 1, \ldots, 6$ to leader's velocity q_{2d}.

TABLE VI
INITIAL VALUES OF $q_{i21}, q_{i22}, i = 1, \ldots, 6$

i	1	2	3	4	5	6
q_{i21}	4.3	2.8	-3.2	-4.6	-1.7	2.6
q_{i22}	0	2.7	1.8	-2.1	-3.2	-4.4

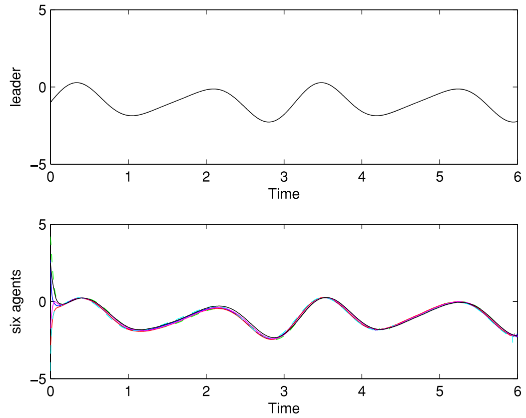


Fig. 6. Velocity trajectory of the first joint.

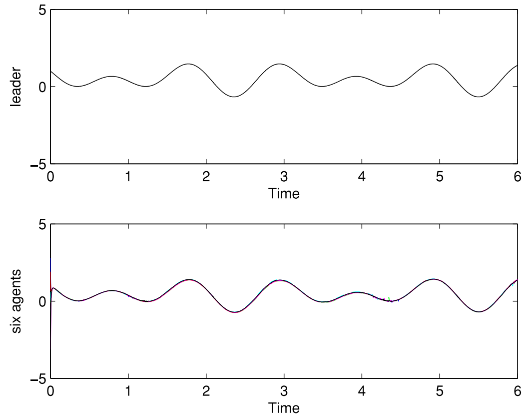


Fig. 7. Velocity trajectory of the second joint.

The consensus controller and the adaptive update law are given by (35) and (36), respectively. The correlation coefficients are chosen as $k_{i0} = 400$, $\gamma_i = 1/16$, $\kappa_i = 12$, $\xi_i = 12$, $\sigma_i = 4$, and $\phi_i = 10^{-7}$, $i = 1, \ldots, 6$.

Figs. 6 and 7 display the simulation results by applying the proposed consensus control scheme to the multimanipulator system (40). They show that the velocity states of the multimanipulator system synchronize to leader's velocity, which means that the proposed method can guarantee a good tracking performance.

Remark 6: If the current NN consensus methods are applied to the two examples, for instance, the consensus method of [27], the NN weight matrices must be 36× 2-dimensional, it means that the consensus algorithm needs to update 72 adaptive parameters online for every agent, so the online computation burden will be heavy and the online computation time will be long. However, if the proposed approach is applied to the nonlinear multiagent (32), only an adaptive parameter is updated for each agent, from the practical viewpoint, the running cost will be greatly decreased.

V. CONCLUSION

In this paper, an adaptive leader-following consensus approach is developed. Because the external disturbances and time delays are considered in the nonlinear multiagent systems, the proposed consensus scheme is more practical and can be applied more widely than the existing agreement methods. In order to finish the control task, the nonlinear uncertainties are counteracted by employing RBFNNs; the state delays are compensated by choosing the appropriate Lyapunov–Krasovskii functional. A remarkable contribution of the work is that the computation burden is alleviated by only updating a small number of adaptive parameters. Finally, the system stability and tracking error convergence are proven by using Lyapunov stability theory. Simulation results further verified the feasibility of the proposed approach.

APPENDIX

Proof: For the case of $e_i(t) \in \Omega^0_{\phi_i}$, choose Lyapunov function candidate as

$$V(t) = V_1(t) + V_2(t) + \frac{1}{2}\sum_{i=1}^{n} \kappa_i^{-1}\tilde{w}_i^2(t) \tag{41}$$

where $\tilde{w}_i(t) = \hat{w}_i(t) - w_i^*$.

Taking the time derivative of (41) along (24) is

$$\begin{aligned}\dot{V} \le& \sum_{i=1}^{n}\left(e_i^T(t)g_i(x_i)u_i + \|e_i(t)\|^2 + e_i^T(t)Q_i(z_i) + \frac{1}{2}\varrho_i^2(x_i)\right)\\ &+ \sum_{i=1}^{n}\kappa_i^{-1}\tilde{w}_i(t)\dot{\hat{w}}_i(t) + \frac{1}{2}\left(\sum_{i=1}^{n}\beta_i^2 + n\alpha^2\right)\end{aligned} \tag{42}$$

where $Q_i(z_i) = f_i(x_i) + (1/2\beta_i^2)e_i(t)q_i^2(x_i)$.

Substituting (34) into (42) yields

$$\begin{aligned}\dot{V}(t) \le& \sum_{i=1}^{n}\Bigg(e_i^T(t)g_i(x_i)u_i + \|e_i(t)\|^2 + e_i^T(t)W_i^*S_i(z_i)\\ &+ e_i^T(t)\varepsilon_i(z_i) + \frac{1}{2}\varrho_i^2(x_i)\Bigg) + \sum_{i=1}^{n}\kappa_i^{-1}\tilde{w}_i(t)\dot{\hat{w}}_i(t)\\ &+ \frac{1}{2}\left(\sum_{i=1}^{n}\beta_i^2 + n\alpha^2\right).\end{aligned} \tag{43}$$

Apply the following facts that

$$\begin{aligned}e_i^T(t)W_i^*S_i(z_i) &\le \xi_i\|e_i(t)\|^2\left\|W_i^*S_i(z_i)\right\|^2 + \frac{1}{4\xi_i}\\ &\le \xi_i w_i^*\|S_i(z_i)\|^2\|e_i(t)\|^2 + \frac{1}{4\xi_i}\end{aligned} \tag{44}$$

where ξ_i is a positive design constant,

$$e_i^T(t)\varepsilon_i(z_i) \le \|e_i(t)\|^2 + \frac{\|\varepsilon_i(z_i)\|^2}{4} \le \|e_i(t)\|^2 + \frac{\delta_i{}^2}{4}, \tag{45}$$

the inequality (43) can be rewritten as

$$\dot V(t) \le \sum_{i=1}^{n}\Bigg(e_i^T(t)g_i(x_i)u_i + 2\|e_i(t)\|^2 + \xi_i w_i^*\|S_i(z_i)\|^2\|e_i(t)\|^2 + \frac{1}{2}\varrho_i^2(x_i)\Bigg) + \sum_{i=1}^{n}\kappa_i^{-1}\tilde w_i(t)\dot{\hat w}_i(t) + \sum_{i=1}^{n}\left(\frac{1}{2}\beta_i^2 + \frac{1}{4\xi_i} + \frac{\delta_i{}^2}{4}\right) + \frac{1}{2}n\alpha^2. \quad (46)$$

Substituting the controller (35) and adaptive law (36) into (46) has

$$\begin{aligned}\dot V(t) \le \sum_{i=1}^{n}\Bigg(&-k_i(t)e_i^T(t)g_i(x_i)e_i(t) - \frac{\xi_i}{\underline{g}_i}\hat w_i(t)\|S_i(z_i)\|^2 e_i^T(t)g_i(x_i)e_i(t)\\ &- \frac{e_i^T(t)g_i(x_i)e_i(t)}{2\underline{g}_i\|e_i(t)\|^2}\varrho_i^2(x_i) + 2\|e_i(t)\|^2 + \xi_i w_i^*\|S_i(z_i)\|^2\|e_i(t)\|^2 + \frac{1}{2}\varrho_i^2(x_i)\Bigg)\\ &+ \sum_{i=1}^{n}\tilde w_i(t)\Big(\xi_i\|S_i(z_i)\|^2\|e_i(t)\|^2 - \sigma_i\hat w_i(t)\Big)\\ &+ \sum_{i=1}^{n}\left(\frac{1}{2}\beta_i^2 + \frac{1}{4\xi_i} + \frac{\delta_i{}^2}{4}\right) + \frac{1}{2}n\alpha^2. \end{aligned} \quad (47)$$

Based on Assumption 5 and Remark 5, (47) can become the following one:

$$\begin{aligned}\dot V(t) &\le \sum_{i=1}^{n}\Big(-k_i(t)e_i^T(t)g_i(x_i)e_i(t) + 2\|e_i(t)\|^2 - \xi_i\hat w_i(t)\|S_i(z_i)\|^2\|e_i(t)\|^2 + \xi_i w_i^*\|S_i(z_i)\|^2\|e_i(t)\|^2\Big)\\ &\quad + \sum_{i=1}^{n}\tilde w_i(t)\Big(\xi_i\|S_i(z_i)\|^2\|e_i(t)\|^2 - \sigma_i\hat w_i(t)\Big) + \sum_{i=1}^{n}\left(\frac{1}{2}\beta_i^2 + \frac{1}{4\xi_i} + \frac{\delta_i{}^2}{4}\right) + \frac{1}{2}n\alpha^2\\ &\le -\sum_{i=1}^{n}\Big(\Big(\underline{g}_i k_i(t) - 2\Big)\|e_i(t)\|^2\Big) - \sum_{i=1}^{n}\sigma_i\tilde w_i(t)\hat w_i(t)\\ &\quad + \sum_{i=1}^{n}\left(\frac{1}{2}\beta_i^2 + \frac{1}{4\xi_i} + \frac{\delta_i{}^2}{4}\right) + \frac{1}{2}n\alpha^2. \end{aligned} \quad (48)$$

Based on the fact that $\tilde w_i(t)\hat w_i(t) = (1/2)\tilde w_i^2(t) + (1/2)\hat w_i^2(t) - (1/2)w_i^{*2}$, the following results can be obtained:

$$-\sigma_i\tilde w_i(t)\hat w_i(t) \le -\frac{1}{2}\sigma_i\tilde w_i^2(t) + \frac{1}{2}\sigma_i w_i^{*2}, \quad i = 1, \ldots, n. \quad (49)$$

The inequality (48) can be rewritten as

$$\dot V(t) \le -\sum_{i=1}^{n}\Big(\underline{g}_i k_i(t) - 2\Big)\|e_i\|^2 - \sum_{i=1}^{n}\sigma_i\tilde w_i^2(t) + \sum_{i=1}^{n}\left(\frac{1}{2}\beta_i^2 + \frac{1}{4\xi_i} + \frac{\delta_i{}^2}{4} + \frac{1}{2}\sigma_i w_i^{*2}\right) + \frac{1}{2}n\alpha^2. \quad (50)$$

From the inequality (37), the inequality (50) can be further rewritten as

$$\begin{aligned}\dot V(t) &\le -\sum_{i=1}^{n}\frac{\gamma_i}{2}\lambda_{\max}(\Delta)\|e_i(t)\|^2 - \sum_{i=1}^{n}\frac{\gamma_i}{2}\int_{t-\tau_{\max}}^{t}\varrho_i^2(x_i(s))ds\\ &\quad - \sum_{i=1}^{n}\sigma_i\tilde w_i^2(t) + \sum_{i=1}^{n}\left(\frac{1}{2}\beta_i^2 + \frac{1}{4\xi_i} + \frac{\delta_i{}^2}{4} + \frac{1}{2}\sigma_i w_i^{*2}\right) + \frac{1}{2}n\alpha^2\\ &\le -\frac{\bar\eta}{2}\sum_{i=1}^{n}\lambda_{\max}(\Delta)\|e_i(t)\|^2 - \frac{\bar\eta}{2}\sum_{i=1}^{n}\int_{t-\tau_{\max}}^{t}\varrho_i^2(x_i(s))ds\\ &\quad - \bar\eta\sum_{i=1}^{n}\kappa_i^{-1}\tilde w_i^2(t) + \bar\theta \end{aligned} \quad (51)$$

where $\bar\eta = \min\{\gamma_1, \ldots, \gamma_n, \sigma_1\kappa_1, \ldots, \sigma_n\kappa_n\}$, $\bar\theta = \sum_{i=1}^{n}((1/2)\beta_i^2 + (1/4\xi_i) + (\delta_i{}^2/4) + (1/2)\sigma_i w_i^{*2}) + (1/2)n\alpha^2$. Obviously, $\bar\theta$ is a positive constant relying on design parameters.

Based on (12) and (31), the inequality (51) can be rewritten as

$$\begin{aligned}\dot V(t) &\le -\bar\eta V_1(t) - \bar\eta V_2(t) - \bar\eta\sum_{i=1}^{n}\kappa_i^{-1}\tilde w_i^2(t) + \bar\theta\\ &\le -\bar\eta V(t) + \bar\theta. \end{aligned} \quad (52)$$

According to Lemma 3, the following inequality can be obtained:

$$V(t) \le \frac{\bar\theta}{\bar\eta} + \left(V(0) - \frac{\bar\theta}{\bar\eta}\right)e^{-\bar\eta t}. \quad (53)$$

It implies that a better consensus tracking performance can be obtained by increasing the gains γ_i and κ_i appropriately.

For the case of $e_i \in \Omega_{\phi_i}$, because ϕ_i is designed to be an arbitrarily small constant, from the inequality (12), it can be directly concluded that the leader-following consensus has arrived. ■

References

[1] W. Ren and R. W. Beard, "A decentralized scheme for spacecraft formation flying via the virtual structure approach," *J. Guid. Control. Dyn.*, vol. 27, no. 1, pp. 73–82, 2004.
[2] J. A. Fax and R. M. Murray, "Information flow and cooperative control of vehicle formations," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1465–1476, Sep. 2004.
[3] T. Eren, P. N. Belhumeur, and A. S. Morse, "Closing ranks in vehicle formations based on rigidity," in *Proc. 41st IEEE Conf. Decis. Control*, vol. 3. Las Vegas, NV, USA, 2002, pp. 2959–2964.
[4] R. Vidal, O. Shakernia, and S. Sastry, "Formation control of nonholonomic mobile robots with omnidirectional visual servoing and motion segmentation," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, vol. 1. Taipei, Taiwan, 2003, pp. 584–589.

[5] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: Algorithms and theory," *IEEE Trans. Autom. Control*, vol. 51, no. 3, pp. 401–420, Mar. 2006.

[6] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1520–1533, Sep. 2004.

[7] W. Ni and D. Z. Cheng, "Leader-following consensus of multi-agent systems under fixed and switching topologies," *Syst. Control Lett.*, vol. 59, nos. 3–4, pp. 209–217, 2010.

[8] Z. K. Li, W. Ren, X. D. Liu, and L. H. Xie, "Distributed consensus of linear multi-agent systems with adaptive dynamic protocols," *Automatica*, vol. 49, no. 7, pp. 1986–1995, 2013.

[9] Q. Song, J. D. Cao, and W. W. Yu, "Second-order leader-following consensus of nonlinear multi-agent systems via pinning control," *Syst. Control Lett.*, vol. 59, no. 9, pp. 553–562, 2010.

[10] Z. K. Li, W. Ren, X. D. Liu, and M. Y. Fu, "Consensus of multi-agent systems with general linear and Lipschitz nonlinear dynamics using distributed adaptive protocols," *IEEE Trans. Autom. Control*, vol. 58, no. 7, pp. 1786–1791, Jul. 2013.

[11] G.-X. Wen, C. L. P. Chen, Y.-J. Liu, and Z. Liu, "Neural-network-based adaptive leader-following consensus control for second-order non-linear multi-agent systems," *IET Control Theory Appl.*, vol. 9, no. 13, pp. 1927–1934, Aug. 2015.

[12] C. L. P. Chen, G.-X. Wen, Y.-J. Liu, and Z. Liu, "Observer-based adaptive backstepping consensus tracking control for high-order nonlinear semi-strict-feedback multiagent systems," *IEEE Trans. Cybern.*, vol. 46, no. 7, pp. 1591–1601, Jul. 2016.

[13] P. Lin and Y. Jia, "Consensus of a class of second-order multi-agent systems with time-delay and jointly-connected topologies," *IEEE Trans. Autom. Control*, vol. 55, no. 3, pp. 778–784, Mar. 2010.

[14] Y. G. Sun and L. Wang, "Consensus problems in networks of agents with double-integrator dynamics and time-varying delays," *Int. J. Control*, vol. 82, no. 10, pp. 1937–1945, 2009.

[15] W. Zhang, J. Liu, D. Zeng, and T. Yang, "Consensus analysis of continuous-time second-order multi-agent systems with nonuniform time-delays and switching topologies," *Asian J. Control*, vol. 15, no. 5, pp. 1516–1523, 2013.

[16] C.-C. Hua, Q.-G. Wang, and X.-P. Guan, "Adaptive fuzzy output-feedback controller design for nonlinear time-delay systems with unknown control direction," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 39, no. 2, pp. 363–374, Apr. 2009.

[17] M. Wang, B. Chen, and P. Shi, "Adaptive neural control for a class of perturbed strict-feedback nonlinear time-delay systems," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 38, no. 3, pp. 721–730, Jun. 2008.

[18] D. W. C. Ho, J. Li, and Y. Niu, "Adaptive neural control for a class of nonlinearly parametric time-delay systems," *IEEE Trans. Neural Netw.*, vol. 16, no. 3, pp. 625–635, May 2005.

[19] B. Xu, C. Yang, and Y. Pan, "Global neural dynamic surface tracking control of strict-feedback systems with application to hypersonic flight vehicle," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 10, pp. 2563–2575, Oct. 2015.

[20] B. Xu, C. Yang, and Z. Shi, "Reinforcement learning output feedback NN control using deterministic learning technique," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 3, pp. 635–641, Mar. 2014.

[21] Y.-J. Liu, G.-X. Wen, and S.-C. Tong, "Direct adaptive NN control for a class of discrete-time nonlinear strict-feedback systems," *Neurocomputing*, vol. 73, nos. 13–15, pp. 2498–2505, 2010.

[22] B. Xu, Z. Shi, C. Yang, and F. Sun, "Composite neural dynamic surface control of a class of uncertain nonlinear systems in strict-feedback form," *IEEE Trans. Cybern.*, vol. 44, no. 12, pp. 2626–2634, Dec. 2014.

[23] B. Xu, Z. Shi, and C. Yang, "Composite fuzzy control of a class of uncertain nonlinear systems with disturbance observer," *Nonlin. Dyn.*, vol. 80, nos. 1–2, pp. 341–351, 2015.

[24] L. Zhang, Z. Ning, and Z. Wang, "Distributed filtering for fuzzy time-delay systems with packet dropouts and redundant channels," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 46, no. 4, pp. 559–572, Apr. 2016.

[25] Y. Li, S. Tong, and T. Li, "Adaptive fuzzy output feedback dynamic surface control of interconnected nonlinear pure-feedback systems," *IEEE Trans. Cybern.*, vol. 45, no. 1, pp. 138–149, Jan. 2015.

[26] Y. Li, S. Tong, and T. Li, "Hybrid fuzzy adaptive output feedback control design for uncertain MIMO nonlinear systems with time-varying delays and input saturation," *IEEE Trans. Fuzzy Syst.*, vol. 24, no. 4, pp. 841–853, Aug. 2016.

[27] C. L. P. Chen, G.-X. Wen, Y.-J. Liu, and F.-Y. Wang, "Adaptive consensus control for a class of nonlinear multiagent time-delay systems using neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 6, pp. 1217–1226, Jun. 2014.

[28] W. S. Chen, X. B. Li, W. Ren, and C. Y. Wen, "Adaptive consensus of multi-agent systems with unknown identical control directions based on a novel Nussbaum-type function," *IEEE Trans. Autom. Control*, vol. 59, no. 7, pp. 1887–1892, Jul. 2014.

[29] T. H. Lee, J. H. Park, D. H. Ji, and H. Y. Jung, "Leader-following consensus problem of heterogeneous multi-agent systems with nonlinear dynamics using fuzzy disturbance observer," *Complexity*, vol. 19, no. 4, pp. 20–31, Mar./Apr. 2014.

[30] S. S. Ge, H. Fan, and T. H. Lee, "Adaptive neural control of nonlinear time-delay systems with unknown virtual control coefficients," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 34, no. 1, pp. 499–516, Feb. 2004.

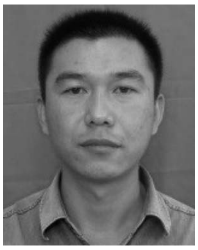
[31] C. W. Wu, *Synchronization in Complex Networks of Nonlinear Dynamical Systems*, vol. 76. World Sci., 2007.

[32] C. L. P. Chen and J. Z. Wan, "A rapid learning and dynamic stepwise updating algorithm for flat neural networks and the application to time-series prediction," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 29, no. 1, pp. 62–72, Feb. 1999.

[33] Q. Song and J. Cao, "On pinning synchronization of directed and undirected complex dynamical networks," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 57, no. 3, pp. 672–680, Mar. 2010.

[34] Z.-G. Hou, L. Cheng, and M. Tan, "Decentralized robust adaptive control for the multiagent system consensus problem using neural networks," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 39, no. 3, pp. 636–647, Jun. 2009.

[35] F. Hong, S. S. Ge, and T. H. Lee, "Practical adaptive neural control of nonlinear systems with unknown time delays," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 35, no. 4, pp. 849–854, Aug. 2005.



Guoxing Wen received the M.S. degree in applied mathematics from the 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 2015 to 2016. He is currently a Lecturer with the Department of Mathematics, Binzhou University, Binzhou, China. His current research interests include adaptive neural network control, optimal control, and multiagent 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, in 1985, and the Ph.D. degree from Purdue University, West Lafayette, IN, USA, in 1988, all 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.

Yan-Jun Liu received the B.S. degree in applied mathematics and the M.S. degree in control theory and control engineering from the Shenyang University of Technology, Shenyang, China, in 2001 and 2004, respectively, and the Ph.D. degree in control theory and control engineering from the Dalian University of Technology, Dalian, China, in 2007.

He is currently a Professor with the College of Science, Liaoning University of Technology, Jinzhou, China. His current research interests include adaptive fuzzy control, nonlinear control, neural network control, and reinforcement learning.

Zhi Liu received the B.S. degree from the Huazhong University of Science and Technology, Wuhan, China, in 1997, the M.S. degree from Hunan University, Changsha, China, in 2000, and the Ph.D. degree from Tsinghua University, Beijing, China, in 2004, all in electrical engineering.

He is currently a Professor with the Department of Automation, Guangdong University of Technology, Guangzhou, China. His current research interests include fuzzy logic systems, neural networks, robotics, and robust control.

报告编号：BZU20180034

文献检索证明

作者姓名：文国兴

作者单位：滨州学院

该作者 2017 年发表在期刊《IEEE TRANSACTIONS ON CYBERNETICS》第 47 卷第 8 期的论文《Neural Network-Based Adaptive Leader-Following Consensus Control for a Class of Nonlinear Multiagent State-Delay Systems》被 SCIE（SCI-EXPANDED）收录。该期刊 2017 年 JCR 的影响因子是 8.803，JCR 分区见附件。在中科院 2016 年 JCR 分区表中位于大类分区 1 区。

检索结果见附件。

证明单位：滨州学院图书馆

证明人：　郭春霞

2018 年 10 月 26 日

关闭 Web of Science 打印
第 1 页 (记录 1 -- 1)
[1]

第 1 条，共 1 条
标题: Neural Network-Based Adaptive Leader-Following Consensus Control for a Class of Nonlinear Multiagent State-Delay Systems
作者: Wen, GX (Wen, Guoxing); Chen, CLP (Chen, C. L. Philip); Liu, YJ (Liu, Yan-Jun); Liu, Z (Liu, Zhi)
来源出版物: IEEE TRANSACTIONS ON CYBERNETICS 卷: 47 期: 8 特刊: SI 页: 2151-2160 DOI: 10.1109/TCYB.2016.2608499 出版年: AUG 2017
Web of Science 核心合集中的 "被引频次": 27
被引频次合计: 27
入藏号: WOS:000405458200032
PubMed ID: 27740504
文献类型: Article
地址: [Wen, Guoxing] Binzhou Univ, Dept Math, Binzhou 256600, Peoples R China.
[Chen, C. L. Philip] Univ Macau, Fac Sci & Technol, Dept Comp & Informat Sci, Macau 256600, Peoples R China.
[Chen, C. L. Philip] UMacau Res Inst, Zhuhai 519080, Peoples R China.
[Liu, Yan-Jun] Liaoning Univ Technol, Coll Sci, Jinzhou 121001, Peoples R China.
[Liu, Zhi] Guangdong Univ Technol, Dept Automat, Guangzhou 510006, Guangdong, Peoples R China.
通讯作者地址: Wen, GX (通讯作者)， Binzhou Univ, Dept Math, Binzhou 256600, Peoples R China.
电子邮件地址: gxwen@live.cn; Philip.Chen@ieee.org; liuyanjun@live.com; lz@gdut.edu.cn
ISSN: 2168-2267
eISSN: 2168-2275
来源出版物页码计数: 10

关闭 Web of Science 打印
第 1 页 (记录 1 -- 1)
[1]

IEEE TRANSACTIONS ON CYBERNETICS

impact factor

8.803 8.775
2017 5 年

JCR®类别	类别中的排序	JCR分区
AUTOMATION & CONTROL SYSTEMS	1/61	Q1
COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE	3/132	Q1
COMPUTER SCIENCE, CYBERNETICS	1/22	Q1

数据来自第 2017 版 Journal Citation Reports

出版商
IEEE-INST ELECTRICAL ELECTRONICS ENGINEERS INC, 445 HOES LANE, PISCATAWAY, NJ 08855-4141 USA

ISSN: 2168-2267
eISSN: 2168-2275

研究领域
Automation & Control Systems
Computer Science

关闭窗口

Artificial Potential-Based Adaptive H_∞ Synchronized Tracking Control for Accommodation Vessel

Guoxing Wen, Shuzhi Sam Ge, *Fellow, IEEE*, Fangwen Tu, and Yoo Sang Choo

***Abstract*—Combining with artificial potential field and robust H_∞ methods, the neural network (NN)-based adaptive synchronized tracking control is proposed for accommodation vessel (AV). The control task is to drive AV synchronous tracking floating production storage and offloading (FPSO). For finishing the task, NN is employed to approximate the unknown nonlinear dynamics of AV; H_∞ method is to guarantee the system states of AV robust to exogenous disturbances; artificial potential method aims to produce the attractive and repulsive forces to assist AV maintaining desired distance with FPSO so that the gangway connecting both AV and FPSO is operated smoothly. Finally, it is proven that the proposed control scheme can guarantee that all error signals of the tracking control are Semi-Globally Uniformly Ultimately Bounded (SGUUB) and AV can synchronously track FPSO to desired accuracy. The simulation results further demonstrate the effectiveness of the proposed method.**

***Index Terms*—Accommodation vessel (AV), artificial potential, neural network (NN), robust H^∞ control, synchronized tracking control.**

I. Introduction

With the increasing demand for exploration and exploitation of offshore oil and gas, more and more offshore operations have to take place in deeper water area. In order to ensure smooth operation for such offshore work, floating production storage and offloading (FPSO) shown in Fig. 1 as working platform always requires the accompany of accommodation vessels (AVs), which are used for providing the space for logistic support and opened deck. Since personnel transportation and equipment transfer between AV and FPSO are achieved by the gangway shown in Fig. 1, AV must be controlled to synchronously track FPSO for finishing these operations smoothly.

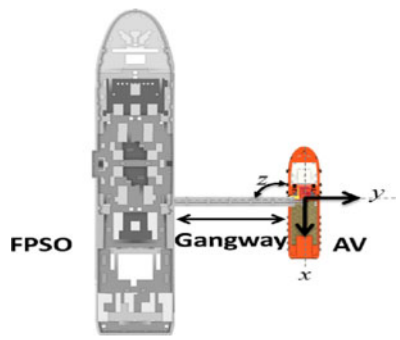


Fig. 1. AV-FPSO system.

Since neural networks (NNs) and fuzzy logic systems (FLSs) have universal approximation ability, which can approximate any smooth nonlinear function to desired accuracy, they have become powerful tools in adaptive nonlinear control. In the past years, many remarkable control schemes using NN or FLS are suggested and received considerable attention, such as [1], [2]. In [1], Rastovic proposes an adaptive recurrent NN synchronization of H-mode and edge-localized mode. Deterministic part of the plasma behavior should be synchronized with stochastic part by introducing stochastic artificial NN. In [2], combined with fuzzy logic and Vlasov–Poisson–Fokker–Planck equations, stability of the tokamak plasma behavior is investigated by the scaling method and Lyapunov functional. Recently, several NN or FLS approximator-based adaptive control methods concerning surface vessel are reported [3]–[8]. In [3], a new NN control is applied to surface vessel control. A significant benefit of the method is that control quality is improved by applying a velocity term besides position term. In [4]–[7], based on NN approximation and backstepping technique, several full-state feedback controls are addressed to tackle system uncertainty problem. According to the Lyapunov stability theory, it is proven that these proposed control methods can guarantee control objective to be achieved. In [8], FLS is applied to dynamic positioning of

Manuscript received May 31, 2016; revised October 18, 2016, December 1, 2016, and January 4, 2017; accepted January 9, 2017. Date of publication March 2, 2017; date of current version June 9, 2017. This work was supported in part by the Doctoral Scientific Research Staring Fund of Binzhou University under Grant 2016Y14 and in part by Keppel-NUS Collaboration Lab under Singapore National Research Foundation under Grant R-261-507-004-281.

G. Wen is with the Department of Mathematics, Binzhou University, Binzhou 256600, China, and also with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore 117576 (e-mail: gxwen@live.cn).

S. S. Ge is with the Department of Electrical and Computer Engineering and the Social Robotics Lab, Interactive Digital Media Institute (IDMI), National University of Singapore, Singapore 117576 (e-mail: samge@nus.edu.sg).

F. Tu is with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore 117576 (e-mail: fangwen_tu@hotmail.com).

Y. S. Choo is with the Centre for Offshore Research and Engineering, National University of Singapore, Singapore 117576 (e-mail: ceecys@nus.edu.sg).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIE.2017.2677330

drill vessel. The proposed fuzzy controller is very simple and does not require the mathematical model of complicated nonlinear system. Finally, the effectiveness of the fuzzy controller are demonstrated by a numerical time-domain simulation.

However, the above control approaches are applied to surface vessel with less regard for the improvement of system robustness. In real marine environment, there exist multiple disturbances, such as wind, wave, swell, and current [9], it is very necessary to consider system robustness in control design. Usually, in order to obtain good system robustness, H_∞ control strategy is naturally considered. Basic idea of the control strategy is to design a control law for dynamic system so that the gain of mapping from exogenous input to measurable output is minimized or is no larger than a certain prescribed level. In the past few decades, NN- or FLS-based H_∞ control has attracted increasing attentions, and many remarkable research results have been reported, for example, [10]–[14]. In [10], robust nonlinear control approach and direct adaptive NN technique are integrated together to construct a new robust learning controller for simultaneous position and force control of uncertain constrained manipulator. In [11] and [12], fuzzy-based H_∞ technique is applied to uncertain nonlinear system. In [13] and [14], the robust H_∞ controls are addressed for the nonlinear stochastic systems.

In the last decades, artificial potential field methods have been extensively investigated and widely applied due to its simplicity and effectiveness. Artificial potential field methods are to fill potential field to workspace so that gradient acting can attract toward the global minimum and repel from the local maxima. Its wide applications can be found in network topology control, robot navigation control, formation control, etc. [15]–[18]. In [15], potential field-based approach is employed to solve the problem of deploying a mobile sensor network in an unknown environment. In [16], artificial potential field method is proposed to deal with unique real-time obstacle avoidance problem for manipulators and mobile robots. In [17], by encoding freespace and goal information to a special artificial potential function, Rimon and Koditschek present a new methodology for exact robot motion planning and control. In [18], artificial potential function and robust control technique are combined for constructing decentralized multiagent formation control scheme.

Motivated by the above discussion, for AV-FPSO system control, it not only requires AV to track FPSO synchronously, but also must guarantee the distance between AV and FPSO in safe range so that the gangway is operated smoothly. Nevertheless, the existing methods, such as [3]–[8], are not specially designed for AV-FPSO systems because these control algorithms do not consider the requirement of operating gangway. The challenging problem is addressed in the paper, the main contributions are summarized in the following.

1) Most previous surface vessel control methods are constructed based on backstepping technique [5]–[7]. Since virtual controller is required, these backstepping-based surface vessel control can only guarantee the position tracking, and it is difficult to ensure the velocity consensus. Because the proposed AV controller is designed to contain both position and velocity control terms, it can steer AV to synchronously track to FPSO.
2) In order to ensure the gangway operating smoothly, artificial potential field method is applied to the proposed synchronized tracking control. Therefore, the risk of damaging the gangway is significantly reduced.
3) By applying the H_∞ control strategy, the good system robustness is guaranteed.

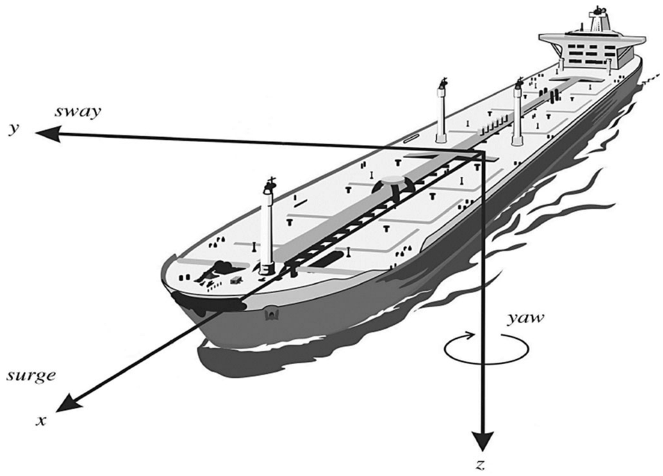


Fig. 2. Three horizontal degrees-of-freedom of surface vessel.

Finally, a simulation is carried out on a scale-down replica of AV to further demonstrate the effectiveness of the proposed scheme.

II. Problem Formulation and Preliminaries

A. Problem Formulation

Consider the surface vessel modeled by the following dynamic, which is depicted in three degree-of-freedom that are surge, sway, and yaw, respectively (shown in Fig. 2) [19]

$$\begin{aligned}\dot{\eta}(t) &= J(\eta)\, v(t),\\ M\dot{v}(t) &= -C(v)\,v(t) - D(v)v(t) - g(\eta(t)) - \Delta(t) + \tau(t)\end{aligned} \tag{1}$$

where $\eta(t) = [x(t), y(t), z(t)]^T \in R^3$ is the state vector, of which $x(t), y(t)$, and $z(t)$ are the position and head states, respectively; $v(t) = [v_x(t), v_y(t), v_z(t)]^T \in R^3$ is the velocity vector, of which $v_x(t), v_y(t), v_z(t)$ are the surge, sway, yaw velocities, respectively;

$$J(\eta) = \begin{bmatrix}\cos(z) & -\sin(z) & 0\\ \sin(z) & \cos(z) & 0\\ 0 & 0 & 1\end{bmatrix} \in SO(3),$$

i.e., $J^{-1}(\eta) = J^T(\eta)$, is the rotation matrix for coordinate transforming between the vessel-fixed and earth-fixed frames; $g(\eta) \in R^3$ is the restoring force vector in the presence of gravity and buoyancy; $\Delta(t) \in R^3$ is the external disturbances, which has the property of $\Delta(t) \in L_2[0, t_p]$, where $\forall t_p \in [0, \infty)$ is the operating time for the system; $\tau(t) \in R^3$ is the control input.

Assumption 1: The position state $\eta(t)$ and velocity state $v(t)$ are measurable without any noises, and they can reflect real states of the vessel.

Let $\{X_{(\cdot)}, Y_{(\cdot)}, N_{(\cdot)}\}$ denote the hydrodynamic parameters [20] and m denote the mass of AV and x_g denote the x coordinate value of gravity center in vessel-fixed frame. Then, all terms of AV dynamic model (1) are detailed in the following.

$M = M^T$ is the system inertia matrix, and is specified as

$$M = \begin{bmatrix} m_{11} & 0 & 0 \\ 0 & m_{22} & m_{23} \\ 0 & m_{32} & m_{33} \end{bmatrix} > 0$$

where $m_{11} = m - X_{\dot{x}}$, $m_{22} = m - Y_{\dot{y}}, m_{33} = I_z - N_{\dot{z}}$, $m_{23} = mx_g - Y_{\dot{z}}, m_{32} = mx_g - N_{\dot{y}}$, and $Y_{\dot{z}} = N_{\dot{y}}$.

$C(v) = -C^T(v)$ is the Coriolis and centripetal matrix, and is specified as

$$C(v) = \begin{bmatrix} 0 & 0 & c_{13} \\ 0 & 0 & c_{23} \\ -c_{13} & -c_{23} & 0 \end{bmatrix}$$

where $c_{13} = -m_{22}v_y - \frac{1}{2}(m_{23} + m_{32})v_z, c_{23} = m_{11}v_x$.

$D(v)$ is the damping matrix specified as

$$D(v) = \begin{bmatrix} d_{11} & 0 & 0 \\ 0 & d_{22} & d_{23} \\ 0 & d_{32} & d_{33} \end{bmatrix}$$

where $d_{11}(v) = -X_x - X_{|x|x}|v_x| - X_{xxx}v_x^2$, $d_{22}(v) = -Y_y - Y_{|y|y}|v_y| - Y_{|z|y}|v_z|$, $d_{33}(v) = -N_z - N_{|y|z}|v_y| - N_{|z|z}|v_z|$, $d_{23}(v) = -Y_z - Y_{|y|z}|v_y| - Y_{|z|z}|v_z|$, $d_{32}(v) = -N_y - N_{|y|y}|v_y| - Y_{|z|y}|v_z|$.

Define $\nu(t) = J(\eta)v(t)$, then the dynamic model (1) can be rewritten as

$$\begin{aligned} \dot{\eta}(t) &= \nu(t), \\ \dot{\nu}(t) &= C_g(\eta, \nu)\nu(t) + D_g(\eta, \nu)\nu(t) \\ &\quad + g_g(\eta(t)) + \Delta_g(t) + \tau_g(t) \end{aligned} \tag{2}$$

where $C_g(\eta, \nu) = -J(\eta)M^{-1}C\left(J^{-1}(\eta)\nu\right)J^{-1}(\eta)$, $D_g(\eta, \nu) = \dot{J}(\eta)J^{-1}(\eta) - J(\eta)M^{-1}D(J^{-1}(\eta)\nu)J^{-1}(\eta)$, $g_g(\eta) = -J(\eta)M^{-1}g(\eta)$, $\Delta_g(t) = -J(\eta)M^{-1}\Delta(t) \in L_2[0, t_p]$, $\tau_g(t) = J(\eta)M^{-1}\tau(t)$.

Let $\eta_r(t)$, $\dot{\eta}_r(t)$, and $\ddot{\eta}_r(t)$ denote the position, velocity, and acceleration of FPSO, where $\eta_r(t) = [x_r(t), y_r(t), z_r(t)]^T \in R^3$. $\eta_r(t)$, $\dot{\eta}_r(t)$, and $\ddot{\eta}_r(t)$ are assumed known and treated as the reference signals followed by AV.

Control objective: Based on the universal approximation property of NN, design an adaptive H_∞ control, such that all error states of the tracking control are SGUUB, while the position and velocity states of AV track FPSO states to desired accuracy. Meanwhile, artificial potential field method is employed to assist AV keeping the desired distance with FPSO for operating the gangway smoothly.

Remark 1: In order to achieve the control objective, the control protocol is designed for the dynamic model (2), so the desired controller is first obtained in earth-fixed frame. Then, the controller for the original dynamic model (1) can be obtained by left multiplying the matrix $MJ^{-1}(z)$.

Lemma 1: ([21]) Let $V(t) \in R$ be a continuous positive function, and its initial value, $V(0)$, is bounded. If $\dot{V}(t) \leq -\beta V(t) + \alpha$ is satisfied, where β and α are positive constants, then the following inequality is held:

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

B. NNs and Function Approximation

It has been proven that radial basis function neural network (RBFNN) has excellent approximation and learning abilities. A continuous nonlinear function $\varphi(z) : R^n \to R^m$ defined on a compact set Ω_z can be approximated by RBFNN in the following form:

$$\varphi_{\rm NN}(z) = W^T S(z) \tag{3}$$

where $W \in R^{p\times m}$ is the adjustable weight matrix, and p is the neuron number; $S(z) = [s_1(z), \ldots, s_p(z)]^T$ is the basis function vector, $s_i(z) = \exp\left[-(z - \mu_i)^T(z - \mu_i)/\phi_i^2\right], i = 1, 2, \ldots, p$, $\mu_i = [\mu_{i1}, \mu_{i2}, \ldots, \mu_{in}]^T$ is the center of the receptive field, ϕ_i is the width of the Gaussian function, $z \in \Omega_z \subset R^n$ is the input vector.

Based on the approximation (3), the continuous vector function $\varphi(z)$ can be re-expressed as

$$\varphi(z) = W^{*T}S(z) + \varepsilon(z) \tag{4}$$

where W^* is the ideal weight; $\varepsilon(z) \in R^m$ is the approximation error, and satisfies $\|\varepsilon(z)\| \leq \delta$, where δ is a positive constant.

The NN approximation error indicates the minimum possible deviation between the optimal approximator $W^{*T}S(z)$ and the unknown function $\varphi(z)$. The ideal NN weight matrix W^* is defined in the following:

$$W^* := \arg\min_{W\in R^{p\times m}}\left\{\sup_{z\in\Omega_z}\|\varphi(z) - WS(z)\|\right\}. \tag{5}$$

In fact, W^* is an "artificial" quantity only for analysis purposes, and it needs to be estimated for the control design [22].

It has been demonstrated that the NN approximation can arrive any desired accuracy if the NN node number is large enough [22]. It implies that $\|\varepsilon(z)\|$ can be reduced to desired smallness if p is sufficiently large.

C. Artificial Potentials Functions

In this paper, artificial potential field methods are employed for controlling AV to keep safe distance with FPSO so that the smooth gangway operation is obtained. The artificial potential is composed of both attractive and repulsive potentials. The corresponding attractive and repulsive forces are produced along with negative gradient direction of the attractive and repulsive potential fields.

Define the relative distance variable $d(t)$ between AV and FPSO as

$$d(t) = \eta(t) - \eta_r(t). \tag{6}$$

The potential function based on the relative position variable is defined as follows.

Definition 1 ([23], [24]): The potential function $P(d(t)) \in R$ is a nonnegative, differentiable, radially unbounded function such that

$P(d(t)) \to \infty$ when $\|d(t)\| \to \infty$;

$P(d(t)) \to \infty$ when $\|d(t)\| \to 0$

$P(d(t))$ attains its unique minimum when $d(t)$ is located at the ideal distance.

The total potential function $P(d(t))$ is designed as

$$P(d(t)) = P_a(d(t)) + P_r(d(t)) \tag{7}$$

where $P_a(d(t))$ and $P_r(d(t))$ denote the attractive and repulsive potential functions, respectively.

The attractive and repulsive forces are derived from the negative gradient of the attractive and repulsive functions, respectively. Then, the total potential force is given as follows:

$$\sigma(d) = \sigma_a(d) + \sigma_r(d) = -\nabla P_a(d) - \nabla P_r(d) \tag{8}$$

where $\sigma_a(d) = -\nabla P_a(d)$ is the attractive force; $\sigma_r(d) = -\nabla P_r(d)$ is the repulsive force; ∇ denotes the gradient corresponding to the distance vector $d(t)$.

The desired distance d_0 between AV and FPSO (the length of the gangway) is designed to be the equilibrium point between the attractive and repulsive forces, i.e., $\sigma_a = -\sigma_r$ when $\|d(t)\| = d_0$. In order to steer AV to keep the desired distance with FPSO, it is requested that the attractive force is bigger than the repulsive force, i.e., $\sigma_a > -\sigma_r$, when $\|d(t)\| > d_0$ implied that AV is moving away from FPSO; the attractive force is smaller than the repulsive force, i.e., $\sigma_a < -\sigma_r$, when $\|d(t)\| < d_0$ implied that AV is closing to FPSO. By integrating the artificial potentials into the synchronized tracking control design, the distance between AV and FPSO can be kept in the desired range.

III. Main Results

Define the tracking error vectors as $e_\eta(t) = \eta(t) - \eta_r(t) - c(t)$, $e_\nu(t) = \nu(t) - \dot{\eta}_r(t) - \dot{c}(t)$, where $c(t)$ is the desired relative distance variable between AV and FPSO, and $\dot{c}(t)$ is its derivative. Then, the error dynamics can be obtained from (2) as follows:

$$\dot{e}_\eta(t) = e_\nu(t), \dot{e}_\nu(t) = f(z(t)) + \Delta_g(t) + \tau_g(t) \tag{9}$$

where $f(z(t)) = C_g(\eta,\nu)\nu(t) + D_g(\eta,\nu)\nu(t) + g_g(\eta(t)) - \ddot{\eta}_r(t) - \ddot{c}(t) \in R^3$, $z(t) = [\eta^T(t), \eta_r^T(t), \nu^T(t), \dot{\eta}_r^T(t), \ddot{\eta}_r^T(t)]^T \in \Omega, \Omega \subset R^{15}$ is a compact set.

A. Robust H_∞ Performance [25], [26]

The output $e(t)$ and disturbance $\omega(t) \in L_2[0, t_p]$ of error dynamic system (9) satisfy the following dissipation inequality:

$$\int_0^{t_p} e^T(t)e(t)dt \le \rho \int_0^{t_p} \omega^T(t)\omega(t)dt + V(0) \tag{10}$$

where ρ is a positive constant, $V(t)$ is a positive semi-definite function, and $\omega(t)$ will be specified later.

Remark 2: The robust H_∞ performance means to attenuate the influence coming from the disturbance input $\omega(t)$ to the tracking error $e(t)$ on a desired level. If the system energy function $V(t)$ starts with zero initial value, i.e., $V(0) = 0$, then the H_∞ performance (10) can be rewritten as $\sup_{\omega \in L_2[0,t_p]} \frac{\|e(t)\|}{\|\omega(t)\|} \le \rho$, which implies that the gain between $e(t)$ and $\omega(t)$ must be equal or less than ρ. Thus, by satisfying the H_∞ control performance (10), it can guarantee the system output to be robust to exogenous disturbances.

Because the nonlinear function $f(z)$ is completely unknown, it cannot be used to the controller design directly. In order to obtain the available controller, RBFNN is employed to approximate the unknown function in the following form:

$$f(z(t)) = W^{*T}S(z(t)) + \varepsilon(z(t)) \tag{11}$$

where $W^* \in R^{p\times 3}$ is the ideal weight matrix, of which p is the neuron number, $S(z(t)) \in R^p$ are the basis function vector, $\varepsilon \in R^3$ is the approximation error to satisfy $\|\varepsilon(t)\| \le \delta$, where δ is a positive constant.

The ideal NN weight W^* is unknown and is given only for analysis purposes, so it needs to be estimated for controller design. Let $\hat{W} \in R^{p\times 3}$ denote the estimation of W^*, then the adaptive controller is constructed in the following:

$$\tau_g(t) = -k_1 e_\eta(t) - k_2 e_\nu(t) - \hat{W}^T(t)S(z) + \sigma(d) \tag{12}$$

where k_1, k_2 are the positive design constants.

The adaptive law for the NN weight matrix is designed as

$$\dot{\hat{W}}(t) = \Gamma\left(S(z)(e_\eta(t) + 2e_\nu(t))^T - \hat{W}(t)\right) \tag{13}$$

where $\Gamma \in R^{p\times p}$ is the positive definite gain matrix.

Based on the controller $\tau_g(t)$ depicted in an earth-fixed frame, the practical controller for the dynamic system (1) is obtained as

$$\tau(t) = MJ^{-1}(\eta)\tau_g(t). \tag{14}$$

The main conclusion can be summarized by the following theorem.

Theorem 1: Consider the AV dynamic modeled by (1) with bounded initial condition, if the design parameters k_1 and k_2 satisfy the following conditions:

$$k_1 \ge \frac{k_2}{2} + 3, k_2 \ge \frac{8}{3} \tag{15}$$

then the proposed adaptive controller (14) can realize the control objective, i.e., all error signals are SGUUB and the distance between AV and FPSO is maintained in safe range by the assistance of artificial potentials, meanwhile, good system robustness is guaranteed by satisfying the H_∞ performance index (10).

Proof: Choose the Lyapunov function candidate as

$$V(t) = \left(k_1 - \frac{1}{2}\right)e_\eta^T(t)e_\eta(t) + \frac{1}{2}e_\nu^T(t)e_\nu(t) + \frac{1}{2}(e_\eta(t) + e_\nu(t))^T(e_\eta(t) + e_\nu(t)) + \frac{1}{2}Tr\left(\tilde{W}^T(t)\Gamma^{-1}\tilde{W}(t)\right) \tag{16}$$

where $\tilde{W}(t) = \hat{W}(t) - W^*$.

Taking the time derivative of $V(t)$ along with (9) and (13), the following result can be obtained:

$$\begin{aligned}\dot V(t) &= (2k_1-1)\,e_\eta^T(t)\,\dot e_\eta(t) + e_\nu^T(t)\,\dot e_\nu(t) + (e_\eta(t)\\&\quad + e_\nu(t))^T(\dot e_\eta(t)+\dot e_\nu(t)) + Tr\left(\tilde W^T(t)\Gamma^{-1}\dot{\hat W}(t)\right)\\&= (2k_1-1)\,e_\eta^T(t)e_\nu(t) + e_\nu^T(f(z)+\Delta_g(t)+\tau_g(t))\\&\quad + (e_\eta(t)+e_\nu(t))^T(e_\nu(t)+f(z)+\Delta_g(t)+\tau_g(t))\\&\quad + Tr\left(\tilde W^T(t)\left(S(z)(e_\eta(t)+2e_\nu(t))^T-\hat W(t)\right)\right).\end{aligned}\tag{17}$$

After several simple manipulations, (17) can be rewritten as

$$\begin{aligned}\dot V(t) &= 2k_1e_\eta^T(t)e_\nu(t) + e_\nu^T(t)e_\nu(t) + (e_\eta(t)+2e_\nu(t))^T\\&\quad (f(z)+\Delta_g(t)+\tau_g(t))\\&\quad + Tr\left(\tilde W^T(t)\left(S(z)\,(e_\eta(t)+2e_\nu(t))^T-\hat W(t)\right)\right).\end{aligned}\tag{18}$$

By substituting (11) and (12) into (18), the following result is obtained:

$$\begin{aligned}\dot V(t) &= 2k_1e_\eta^T(t)e_\nu(t) + e_\nu^T(t)e_\nu(t) + (e_\eta(t)+2e_\nu(t))^T\\&\quad \big(W^{*T}S(z)+\varepsilon(z)+\Delta_g(t)-k_1e_\eta(t)-k_2e_\nu(t)\\&\quad -\hat W^T(t)S(z)+\sigma(d)\big)\\&\quad + Tr\left(\tilde W^T(t)\left(S(z(t))\,(e_\eta(t)+2e_\nu(t))^T-\hat W(t)\right)\right)\\&= -k_1e_\eta^T(t)e_\eta(t)-(2k_2-1)e_\nu^T(t)e_\nu(t)-k_2e_\eta^T(t)e_\nu(t)\\&\quad -(e_\eta(t)+2e_\nu(t))^T\tilde W^T(t)S(z)+(e_\eta(t)+2e_\nu(t))^T\\&\quad (\sigma(d)+\Delta_g(t)+\varepsilon(z))\\&\quad + Tr\left(\tilde W^T(t)\left(S(z)\,(e_\eta(t)+2e_\nu(t))^T-\hat W(t)\right)\right).\end{aligned}\tag{19}$$

Using the property of trace operator that $a^Tb=Tr(ab^T)=Tr(ba^T)$, $a,b\in R^n$, the following can be obtained:

$$\begin{aligned}&Tr\left(\tilde W^T(t)S(x)(e_\eta(t)+2e_\nu(t))^T\right)\\&\quad = (e_\eta(t)+2e_\nu(t))^T\tilde W^T(t)S(z).\end{aligned}\tag{20}$$

Substituting (20) into (19), the following result is obtained:

$$\begin{aligned}\dot V(t) &= -k_1e_\eta^T(t)e_\eta(t)-(2k_2-1)e_\nu^T(t)e_\nu(t)-k_2e_\eta^T(t)e_\nu(t)\\&\quad + (e_\eta(t)+2e_\nu(t))^T(\sigma(d)+\Delta_g(t)+\varepsilon(z))-(e_\eta(t)\\&\quad +2e_\nu(t))^T\,\tilde W^T(t)S(z)+(e_\eta+2e_\nu)^T\,\tilde W^T(t)S(z)\\&\quad - Tr\left(\tilde W^T(t)\hat W(t)\right)\\&= -k_1e_\eta^T(t)e_\eta(t)-(2k_2-1)e_\nu^T(t)e_\nu(t)-k_2e_\eta^T(t)e_\nu(t)\\&\quad + e_\eta^T(t)\sigma(d)+2e_\nu^T(t)\sigma(d)+(e_\eta(t)+2e_\nu(t))^T(\Delta_g(t)\\&\quad +\varepsilon(z))-Tr\left(\tilde W^T(t)\hat W(t)\right).\end{aligned}\tag{21}$$

Add and subtract $\frac{k_2}{2}e_\eta^T(t)e_\eta(t)$ and $\frac{k_2}{2}e_\nu^T(t)e_\nu(t)$ to the right-hand side of (21) to yield the following result:

$$\begin{aligned}\dot V(t) &= -\left(k_1-\frac{k_2}{2}\right)e_\eta^T(t)e_\eta(t)-\left(1\frac{1}{2}k_2-1\right)e_\nu^T(t)e_\nu(t)\\&\quad -\frac{k_2}{2}e_\eta^T(t)e_\eta(t)-k_2e_\eta^T(t)e_\nu(t)-\frac{k_2}{2}e_\nu^T(t)e_\nu(t)\\&\quad + e_\eta^T(t)\sigma(d)+2e_\nu^T\sigma(d)+(e_\eta(t)+2e_\nu(t))^T\\&\quad (\Delta_g(t)+\varepsilon(z))-Tr\left(\tilde W^T(t)\hat W(t)\right)\\&= -\left(k_1-\frac{k_2}{2}\right)e_\eta^T(t)e_\eta(t)-\left(1\frac{1}{2}k_2-1\right)e_\nu^T(t)e_\nu(t)\\&\quad -\frac{k_2}{2}(e_\eta(t)+e_\nu(t))^T(e_\eta(t)+e_\nu(t))+e_\eta^T(t)\sigma(d)\\&\quad + 2e_\nu^T(t)\sigma(d)+(e_\eta(t)+2e_\nu(t))^T(\Delta_g(t)+\varepsilon(z))\\&\quad - Tr\left(\tilde W^T(t)\hat W(t)\right).\end{aligned}\tag{22}$$

According to the Cauchy inequality, $\left(\sum_{k=1}^n a_kb_k\right)^2\le\sum_{k=1}^n a_k^2\sum_{k=1}^n b_k^2$, and Young's inequality, $ab\le\frac{a^2}{2}+\frac{b^2}{2}$, the following facts hold:

$$e_\eta^T(t)\sigma(d)\le e_\eta^T(t)e_\eta(t)+\frac{1}{4}\sigma^T(d)\,\sigma(d),\tag{23}$$

$$2e_v^T(t)\sigma(d)\le e_v^T(t)e_v(t)+\sigma^T(d)\,\sigma(d).\tag{24}$$

$$(e_\eta+2e_\nu)^T\Delta_g(t)\le e_\eta^Te_\eta+e_\nu^Te_\nu+1\frac{1}{4}\Delta_g^T(t)\Delta_g(t),\tag{25}$$

$$(e_\eta+2e_\nu)^T\varepsilon(z)\le e_\eta^Te_\eta+e_\nu^Te_\nu+1\frac{1}{4}\varepsilon^T(z)\varepsilon(z).\tag{26}$$

Using (23)–(26), (22) can become the following one:

$$\begin{aligned}\dot V(t) &\le -\left(k_1-\frac{k_2}{2}-3\right)e_\eta^T(t)e_\eta(t)-\left(1\frac{1}{2}k_2-4\right)\\&\quad e_\nu^T(t)e_\nu(t)-\frac{k_2}{2}(e_\eta(t)+e_\nu(t))^T(e_\eta(t)+e_\nu(t))\\&\quad -Tr\left(\tilde W^T(t)\hat W(t)\right)+1\frac{1}{4}\sigma^T(d)\sigma(d)+1\frac{1}{4}\Delta_g^T(t)\\&\quad \Delta_g(t)+1\frac{1}{4}\varepsilon^T(z)\varepsilon(z).\end{aligned}\tag{27}$$

Applying the fact $Tr\left(\tilde W^T(t)\hat W(t)\right)=\frac{1}{2}Tr\left(\tilde W^T(t)\tilde W(t)\right)+\frac{1}{2}Tr\left(\hat W^T(t)\hat W(t)\right)-\frac{1}{2}Tr\left(W^{*T}W^*\right)$ to the above inequality, the following is obtained:

$$\dot{V}(t) \leq -\left(k_1 - \frac{k_2}{2} - 3\right) e_\eta^T(t) e_\eta(t) - \left(1\frac{1}{2}k_2 - 4\right) e_\nu^T(t) e_\nu(t) - \frac{k_2}{2}\left(e_\eta(t) + e_\nu(t)\right)^T \left(e_\eta(t) + e_\nu(t)\right) - \frac{\gamma}{2} Tr\left(\tilde{W}^T(t)\Gamma^{-1}\tilde{W}(t)\right) + \omega^T(t)\omega(t) \tag{28}$$

where $\gamma = \lambda_{\max}(\Gamma)$, $\omega^T(t)\omega(t) = 1\frac{1}{4}\sigma^T(d)\sigma(d) + 1\frac{1}{4}\Delta_g^T(t)\Delta_g(t) + 1\frac{1}{4}\varepsilon^T(z)\varepsilon(z) + \frac{1}{2}Tr\left(W^{*T}W^*\right)$.

Since AV and FPSO are connected by the gangway, the relative distance variable $d(t)$ is limited in a neighborhood of the equilibrium point. According to the definition of artificial potential, it can be concluded that $\sigma(d) \in L_2[0, t_p]$, associated with the facts that $\Delta_g(t) \in L_2[0, t_p]$ and $\varepsilon(z)$ are bounded, the term $\omega^T(t)\omega(t)$ can be bounded by a constant α, i.e., $\|\omega(t)\|^2 \leq \alpha$. Let $\beta = \min\{\frac{k_1 - \frac{k_2}{2} - 3}{k_1 - \frac{1}{2}}, 3k_2 - 8, k_2, \gamma\}$, the inequality (28) can become the following one:

$$\dot{V}(t) \leq -\beta V(t) + \alpha \tag{29}$$

where $\beta > 0$ can be guaranteed when the design constants k_1, k_2 satisfy (15). According to Lemma 1, the following inequality can be obtained:

$$V(t) \leq V(0)e^{-\beta t} + \frac{\alpha}{\beta}\left(1 - e^{-\beta t}\right). \tag{30}$$

The above inequality implies that all error states are SGUUB and the position and velocity of AV can track FPSO states to desired accuracy by choosing suitable design parameters.

In addition, from the inequality (28), there is the following inequality:

$$\dot{V}(t) \leq -\left(k_1 - \frac{k_2}{2} - 3\right) e_\eta^T(t)e_\eta(t) - \left(1\frac{1}{2}k_2 - 4\right) e_\nu^T(t)e_\nu(t) + \omega^T(t)\omega(t). \tag{31}$$

Let $\mu = \min\{k_1 - \frac{k_2}{2} - 3, 1\frac{1}{2}k_2 - 4\}$, the inequality (31) can become the following one:

$$\dot{V}(t) \leq= -\mu e^T(t)e(t) + \omega^T(t)\omega(t). \tag{32}$$

Integrating the inequality (32) from $t = 0$ to $t = t_p$, the following inequality can be obtained:

$$V(t_p) - V(0) \leq -\mu \int_0^{t_p} e^T(t)e(t)dt + \int_0^{t_p} \omega^T\omega dt. \tag{33}$$

Based on the fact $V(t_p) \geq 0$, the inequality (33) can become the following one:

$$\int_0^{t_p} e^T(t)e(t)dt \leq \rho \int_0^{t_p} \omega^T(t)\omega(t)dt + V(0) \tag{34}$$

where $\rho = 1/\mu$.

Finally, H_∞ performance is satisfied, which implies that the proposed control approach can realize the control objective. ■

Remark 3: Most existed research results for surface vessel control are based on backstepping techniques, for example, [6], [27]. Since virtual controllers are required in backstepping control, it is difficult to achieve the velocity consensus. Especially, for high-order system control, the technique even possibly causes the problem of "explosion of complexity" by repeatedly taking the derivative of virtual controllers. The proposed control strategy skips the virtual controller design, hence it can achieve both the position and velocity consensuses by integrating both position and velocity error terms into the controller design.

TABLE I
HYDRODYNAMIC PARAMETERS

I_z	1.7	$Y_{\|z\|z}$	– 2	$N_{\|y\|z}$	– 4.0
x_g	0.04	$Y_{\|y\|y}$	– 36	$N_{\|z\|z}$	– 4
X_x	– 0.72	$Y_{\|y\|z}$	2	$X_{\dot{x}}$	– 2.0
$X_{\|x\|x}$	– 1.3	$Y_{\|z\|y}$	– 3	$Y_{\dot{y}}$	– 10
X_{xxx}	– 5.8	N_y	0.1	$Y_{\dot{z}}$	– 0.0
Y_y	– 0.86	N_z	– 6.0	$N_{\dot{y}}$	– 0.0
Y_z	0.1	$N_{\|y\|y}$	5.0	$N_{\dot{z}}$	– 1.0

IV. SIMULATION EXAMPLES

In order to further demonstrate the effectiveness of the proposed synchronized tracking control, a simulation example is carried out by a scale-down replica of AV. Its mass is $m = 18$ kg, the length is 1.2 m, and the width is 0.3 m. All hydrodynamic parameters of the model ship are shown in Table I. The inertia, centrifugal and Coriolis, and damping matrices can be calculated as follows:

$$M = \begin{bmatrix} 20 & 0 & 0 \\ 0 & 19 & 0.72 \\ 0 & 0.72 & 2.7 \end{bmatrix}, \Delta(t) = \begin{bmatrix} \eta_x^2(t)\cos(1.5t) \\ \eta_y^2(t)\sin(t) \\ 0 \end{bmatrix},$$

$$C = \begin{bmatrix} 0 & 0 & -19v_y - 0.72v_z \\ 0 & 0 & 20v_x \\ 19v_y + 0.72v_z & -20v_x & 0 \end{bmatrix},$$

$$D = \begin{bmatrix} 0.72 + 1.3|v_x| + 5.8v_x^2 & 0 & 0 \\ 0 & 0.86 + 36|v_y| + 3|v_z| & -0.1 - 2|v_y| + 2|v_z| \\ 0 & -0.1 - 5|v_y| + 3|v_z| & 6 + 4|v_y| + 4|v_z| \end{bmatrix}.$$

For simplicity sake, the restoring force vector $g(\eta(t))$ is assumed to be 0, i.e., $g(\eta(t)) = 0$, which can be found in a large number of literatures, for example, [6], [27].

The initial values for the position and velocity states are $[13.5, 0, 0]^T$ and $[0, 0, 0]^T$, respectively. The desired reference signals $\eta_r(t)$ and $\dot{\eta}_r(t)$ are given as $\eta_r(t) = \begin{bmatrix} 12\sin\left(0.2t + \frac{\pi}{2}\right) \\ 12\sin(0.2t) \\ \arcsin(\sin(0.2t)) + \frac{\pi}{2} \end{bmatrix}$ and $\dot{\eta}_r(t) = \begin{bmatrix} 2.4\cos\left(0.2t + \frac{\pi}{2}\right) \\ 2.4\cos(0.2t) \\ 0.2 \end{bmatrix}$. The initial values for the position and velocity states of reference signal are $[12, 0, 0]^T$ and $[2, 0, 0]^T$.

The tracking error vectors are $e_\eta(t) = \eta(t) - \eta_r(t) - c(t)$, $e_\nu(t) = \nu(t) - \dot{\eta}_r(t) - \dot{c}(t)$, respectively, where $c(t) = \left[d_0\sin(0.2t), d_0\sin\left(0.2t + \frac{\pi}{2}\right), 0\right]^T$ is the desired relative

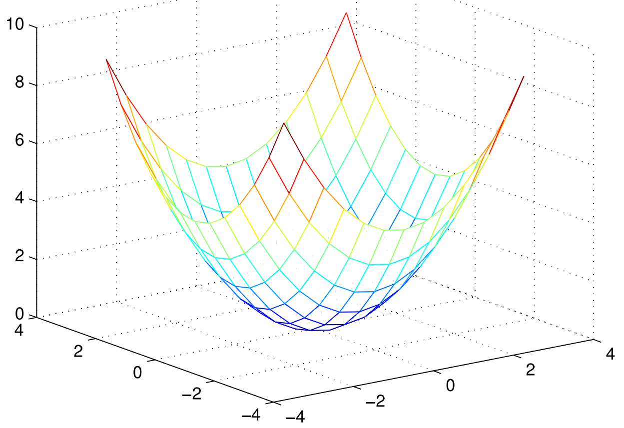


Fig. 3. Attractive potential.

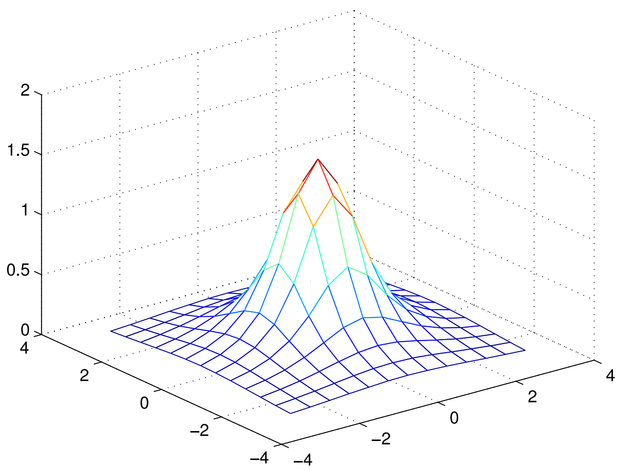


Fig. 4. Repulsive potential.

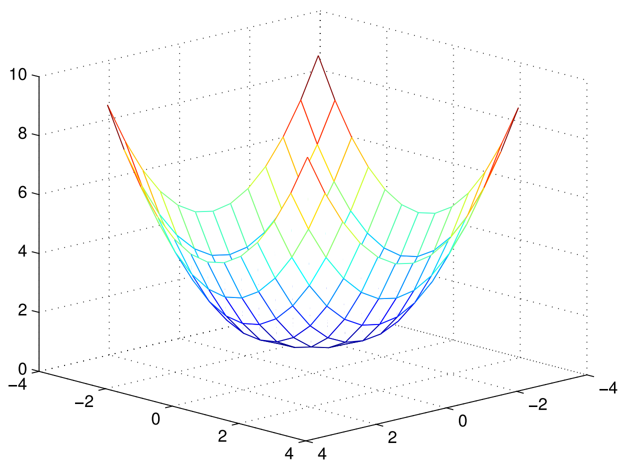


Fig. 5. Total potential.

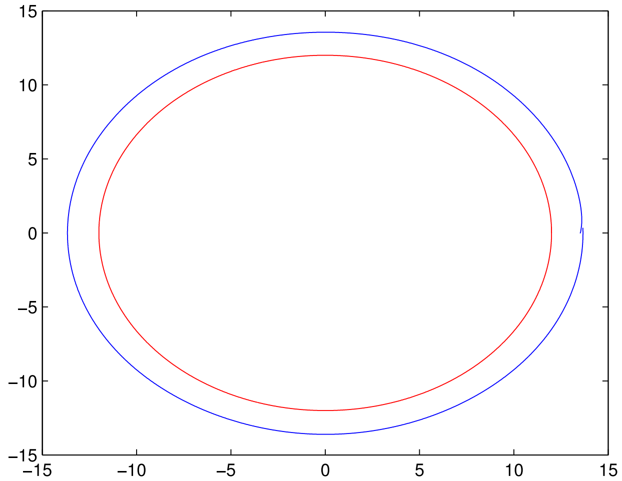


Fig. 6. Trajectories of AV and reference signal.

position variable between AV and FPSO, where $d_0 = 4$ is the length of the gangway.

Remark 4: It should be mentioned that the head state control of AV is explicit because the trajectory of FPSO is a circle. For simplicity, the control of this state is not specified and the performance figure is not displayed.

In this example, RBFNN is chosen to approximate the unknown function. The RBFNN is designed to contain 60 nodes, i.e., $p = 60$. The centers, μ_i, evenly spaced in the range of $\underbrace{[-15,15]\times[-15,15]\times[-3,3]}_{\times 2}\underbrace{[-3,3]\times[-3,3]\times[-3,3]}_{\times 3}$, and the widths are $\phi_i = 2$ for all. The initial conditions for the weight matrix is $W(0) = 0_{3\times 60}$, and the design constants for adaptive law (13) are chosen as $\Gamma = 4$.

The attractive and repulsive potential functions are specified as

$$P_a(d) = \alpha\|e_\eta(t)\|^2$$
$$P_r(d) = \beta \operatorname{arccot}\left(\|e_\eta(t)\|^2\right) \tag{35}$$

where α and β are positive design parameters, which are specified later. Figs. 3 and 4 show the attractive and repulsive potential, respectively. The total potential is shown in Fig. 5.

The corresponding attractive and repulsive forces are expressed by the following equations:

$$\sigma_a(d(t)) = -\nabla P_a(d(t)) = -2\alpha e_\eta(t)$$
$$\sigma_r(d(t)) = -\nabla P_r(d(t)) = \frac{2\beta}{1+\|e_\eta(t)\|^4} e_\eta(t). \tag{36}$$

When α and β satisfy the condition that $\beta = \alpha$, the equilibrium position between the attractive and repulsive forces can be placed at $d(t) = d_0 = 4$. Then, the design constants for the controller (12) are chosen as $k_1 = 120$, $k_2 = 80$, $\alpha = \beta = 100$.

The simulation results are displayed in Figs. 6–8. Fig. 6 shows the position states of AV to track the trajectory of the desired reference. Fig. 7 shows the movement trajectory of AV without the assistance of artificial potentials, and the gangway cannot be run smoothly in the absence of artificial potentials. Fig. 8 shows that the velocity of AV can follow to desired velocity by the proposed control method.

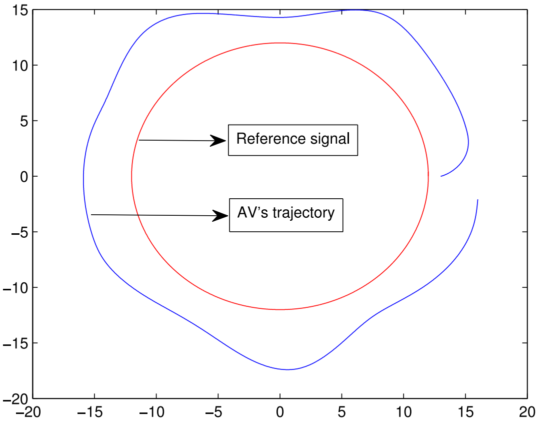


Fig. 7. Trajectories of AV without the assistance of artificial potential.

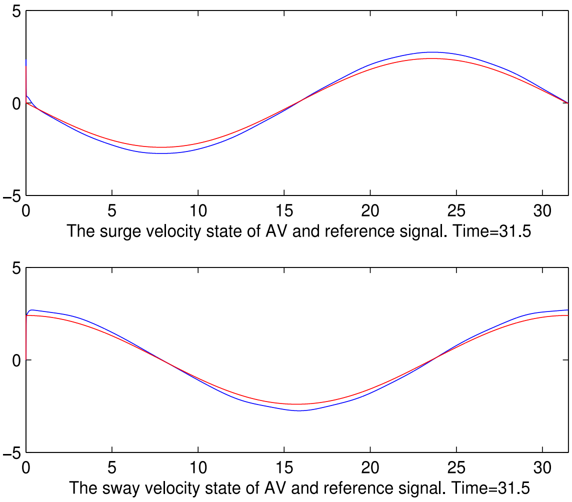


Fig. 8. Velocity vectors of AV and reference.

V. Conclusion

Based on the excellent approximation property of adaptive NN, the proposed robust H_∞ tracking control can be well applied to AV-FPSO systems. Artificial potential field method was employed to assist AV to keep the desired distance with FPSO. Since both position and velocity terms are integrated into the adaptive vessel controller, the proposed control strategy can guarantee that all error signals of the tracking control are SGUUB and AV can synchronously track to FPSO. The simulation example was carried out to further demonstrate the effectiveness of the proposed approach.

Acknowledgment

The conclusions put forward reflect the views of the authors alone, and not necessarily those of the institutions within the Corporate Laboratory.

References

[1] D. Rastovic, "Targeting and synchronization at tokamak with recurrent artificial neural networks," *Neural Comput. Appl.*, vol. 21, no. 5, pp. 1065–1069, 2012.
[2] D. Rastovic, "Fuzzy scaling and stability of tokamaks," *J. Fusion Energy*, vol. 28, no. 1, pp. 101–106, 2009.
[3] D. Li and M. X. Gu, "Dynamic positioning of ships using a planned neural network controller," *J. Ship Res.*, vol. 40, no. 2, pp. 164–171, 1996.
[4] M. Chen, S. S. Ge, and Y. S. Choo, "Neural network tracking control of ocean surface vessels with input saturation," in *Proc. IEEE Int. Conf. Autom. Logist.*, 2009, pp. 85–89.
[5] M. Chen, S. S. Ge, B. V. E. How, and Y. S. Choo, "Robust adaptive position mooring control for marine vessels," *IEEE Trans. Control Syst. Technol.*, vol. 21, no. 2, pp. 395–409, Mar. 2013.
[6] K. P. Tee and S. S. Ge, "Control of fully actuated ocean surface vessels using a class of feedforward approximators," *IEEE Trans. Control Syst. Technol.*, vol. 14, no. 4, pp. 750–756, Jul. 2006.
[7] Z. Zhao, W. He, and S. S. Ge, "Adaptive neural network control of a fully actuated marine surface vessel with multiple output constraints," *IEEE Trans. Control Syst. Technol.*, vol. 22, no. 4, pp. 1536–1543, Jul. 2014.
[8] T.-H. Lee, Y. Cao, and Y.-M. Lin, "Dynamic positioning of drilling vessels with a fuzzy logic controller," *Int. J. Syst. Sci.*, vol. 33, no. 12, pp. 979–993, 2002.
[9] T. I. Fossen, *Handbook of Marine Craft Hydrodynamics and Motion Control*. Hoboken, NJ, USA: Wiley, 2011.
[10] M.-C. Hwang and X. Hu, "A robust position/force learning controller of manipulators via nonlinear H_∞ control and neural networks," *IEEE Trans. Syst., Man Cybern. B, Cybern.*, vol. 30, no. 2, pp. 310–321, Apr. 2000.
[11] W.-Y. Wang, M.-L. Chan, C.-C. J. Hsu, and T.-T. Lee, "H_∞ tracking-based sliding mode control for uncertain nonlinear systems via an adaptive fuzzy-neural approach," *IEEE Trans. Syst. Man Cybern. B, Cybern.*, vol. 32, no. 4, pp. 483–492, Aug. 2002.
[12] Y.-C. Chang, "Adaptive fuzzy-based tracking control for nonlinear SISO systems via VSS and H_∞ approaches," *IEEE Trans. Fuzzy Syst.*, vol. 9, no. 2, pp. 278–292, Apr. 2001.
[13] W. Zhang, B.-S. Chen, and C.-S. Tseng, "Robust H_∞ filtering for nonlinear stochastic systems," *IEEE Trans. Signal Process.*, vol. 53, no. 2, pp. 589–598, Feb. 2005.
[14] W. Zhang and B.-S. Chen, "State feedback H_∞ control for a class of nonlinear stochastic systems," *SIAM J. Control Optim.*, vol. 44, no. 6, pp. 1973–1991, 2006.
[15] A. Howard, M. J. Matarić, and G. S. Sukhatme, "Mobile sensor network deployment using potential fields: A distributed, scalable solution to the area coverage problem," in *Distributed Autonomous Robotic Systems 5*. New York, NY, USA: Springer, 2002, pp. 299–308.
[16] O. Khatib, "Real-time obstacle avoidance for manipulators and mobile robots," *Int. J. Robot. Res.*, vol. 5, no. 1, pp. 90–98, 1986.
[17] E. Rimon and D. E. Koditschek, "Exact robot navigation using artificial potential functions," *IEEE Trans. Robot. Autom.*, vol. 8, no. 5, pp. 501–518, Oct. 1992.
[18] B. Ranjbar-Sahraei, F. Shabaninia, A. Nemati, and S.-D. Stan, "A novel robust decentralized adaptive fuzzy control for swarm formation of multi-agent systems," *IEEE Trans. Ind. Electron.*, vol. 59, no. 8, pp. 3124–3134, Aug. 2012.
[19] T. Fossen, *Marine Control Systems: Guidance, Navigation, and Control of Ships, Rigs and Underwater Vehicles*. Trondheim, Norway: Mar. Cybern., 2012.
[20] A. Caruba, "SURVEYOR," Soc. Naval Architects Marine Eng., Alexandria, VI, USA , 1978.
[21] Z. G. Hou, L. C. Chen, and M. Tan, "Decentralized robust adaptive control for the multiagent system consensus problem using neural networks," *IEEE Trans. Syst., Man Cybern. B, Cybern.*, vol. 39, no. 3, pp. 636–647, Jun. 2009.
[22] S. S. Ge, C. C. Hang, and T. Zhang, "Adaptive neural network control of nonlinear systems by state and output feedback," *IEEE Trans. Syst. Man Cybern. B, Cybern.*, vol. 29, no. 6, pp. 818–828, Dec. 1999.
[23] H. G. Tanner, A. Jadbabaie, and G. J. Pappas,"Stable flocking of mobile agents, part i: Fixed topology," in *Proc 42nd IEEE Conf. Decision Control*, 2003, vol. 2, pp. 2010–2015.
[24] L. Wang, Z. Chen, Z. Liu, and Q. Wang, "Formation control of multi-agent system based on potential function," in *Proc. Asia Simul. Conf. 7th Int. Conf. Syst. Simul. Sci. Comput.*, 2008, pp. 101–105.
[25] B.-S. Chen, C.-H. Lee, and Y.-C. Chang, "H_∞ tracking design of uncertain nonlinear siso systems: Adaptive fuzzy approach," *IEEE Trans. Fuzzy Syst.*, vol. 4, no. 1, pp. 32–43, Feb. 1996.
[26] Y. Yang and C. Zhou, "Adaptive fuzzy H_∞ stabilization for strict-feedback canonical nonlinear systems via backstepping and small-gain approach," *IEEE Trans. Fuzzy Syst.*, vol. 13, no. 1, pp. 104–114, Feb. 2005.
[27] R. Skjetne, T. I. Fossen, and P. V. Kokotović, "Adaptive maneuvering, with experiments, for a model ship in a marine control laboratory," *Automatica*, vol. 41, no. 2, pp. 289–298, 2005.

Authors' photographs and biographies not available at the time of publication.

报告编号：BZU20180035

文献检索证明

作者姓名：文国兴

作者单位：滨州学院

该作者 2017 年发表在期刊《IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS 》第 64 卷第 7 期的论文《Artificial Potential-Based Adaptive H-infinity Synchronized Tracking Control for Accommodation Vessel》被 SCIE（SCI-EXPANDED）收录。该期刊 2017 年 JCR 的影响因子是 7.05，JCR 分区见附件。在中科院 2016 年 JCR 分区表中位于大类分区 1 区。

检索结果见附件。

证明单位：滨州学院图书馆

证明人： 郭春霞

2018 年 10 月 26 日

关闭

Web of Science
第 1 页 (记录 1 -- 1)
◀[1]▶

打印

第 1 条，共 1 条

标题: Artificial Potential-Based Adaptive H-infinity Synchronized Tracking Control for Accommodation Vessel

作者: Wen, GX (Wen, Guoxing); Ge, SS (Ge, Shuzhi Sam); Tu, FW (Tu, Fangwen); Choo, YS (Choo, Yoo Sang)

来源出版物: IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS 卷: 64 期: 7 页: 5640-5647 DOI: 10.1109/TIE.2017.2677330 出版年: JUL 2017

Web of Science 核心合集中的 "被引频次": 4

被引频次合计: 4

入藏号: WOS:000403011200039

文献类型: Article

地址: [Wen, Guoxing] Binzhou Univ, Dept Math, Binzhou 256600, Peoples R China.
[Wen, Guoxing; Ge, Shuzhi Sam; Tu, Fangwen] Natl Univ Singapore, Dept Elect & Comp Engn, Singapore 117576, Singapore.
[Ge, Shuzhi Sam] Natl Univ Singapore, Social Robot Lab, IDMI, Singapore 117576, Singapore.
[Choo, Yoo Sang] Natl Univ Singapore, Ctr Offshore Res & Engn, Singapore 117576, Singapore.
通讯作者地址: Wen, GX (通讯作者)， Binzhou Univ, Dept Math, Binzhou 256600, Peoples R China.
电子邮件地址: gxwen@live.cn; samge@nus.edu.sg; fangwen_tu@hotmail.com; ceecys@nus.edu.sg

ISSN: 0278-0046
eISSN: 1557-9948
来源出版物页码计数: 8

关闭

Web of Science
第 1 页 (记录 1 -- 1)
◀[1]▶

打印

IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS

impact factor

7.05	7.623
2017	5 年

JCR®类别	类别中的排序	JCR分区
AUTOMATION & CONTROL SYSTEMS	2/61	Q1
ENGINEERING, ELECTRICAL & ELECTRONIC	13/260	Q1
INSTRUMENTS & INSTRUMENTATION	1/61	Q1

数据来自第 2017 版 Journal Citation Reports

出版商
IEEE-INST ELECTRICAL ELECTRONICS ENGINEERS INC, 445 HOES LANE, PISCATAWAY, NJ 08855-4141 USA

ISSN: 0278-0046
eISSN: 1557-9948

研究领域
Automation & Control Systems
Engineering
Instruments & Instrumentation

关闭窗口

附件 3

申报 2019 年度省有突出贡献的中青年专家推荐人选个人承诺书（样本）

本人自愿申请，经所在单位和主管部门（单位）同意，申报 2019 年度山东省有突出贡献的中青年专家，并郑重承诺:

本人所填报提交的个人信息、材料内容均真实、准确、有效，并与本人实际情况完全相符。本人未入选过上层次人才工程，不在其他省级重点人才工程管理期内（或已征得省主管部门审核同意)，不存在多头申报、重复申报等行为。

如本人入选，将自觉履行省有关规定，5 年管理期内不申报其他人才工程或类别，对因提供申报材料不实或违反有关规定引起的后果，本人愿承担相关责任。

承诺人（签字）: 文国兴

身份证号码: 37230119770215075X

联系电话: 0543 3191176

工作单位（盖章）: 滨州学院

2019年 7月 2日

承诺书.jpg

滨州学院人事处

证　明

兹证明文国兴，男，身份证号：37230119770215075X，为我校理学院在职教师，为正式在编人员，特此证明。

滨州学院人事处

2019 年 7 月 03 日

教师证明.jpg

滨州学院人事处

关于文国兴同志工作经历说明

文国兴，男，身份证号：37230119770215075X，教授职称，其本人工作经历如下：

1997.09—2011.09　滨州市梁才乡教委教师

2016.01—至今　　滨州学院理学院教师

（其中，2015.09—2016.09　新加坡国立大学从事博士后工作

2018.06-2018.09　澳门大学科技学院从事博士后工作　）

特此证明。

滨州学院人事处

2019 年 7 月 3 日

滨州学院人事处

关于文国兴同志工作经历说明

文国兴，男，身份证号：37230119770215075X，教授职称，其本人工作经历如下：

1997.09—2011.09　滨州市梁才乡教委教师

2016.01—至今　　滨州学院理学院教师

（其中，2015.09—2016.09　新加坡国立大学从事博士后工作

2018.06-2018.09 澳门大学科技学院从事博士后工作 ）

特此证明。

滨州学院人事处

2019 年 7 月 3 日

留 学 回 国 人 员 证 明

（ 2016 新加坡 ） 教(文) 证字 2108 号

兹证明 文国兴 （男□ v、女□，护照号码 E00098555 ）系我国

在 新加坡 国 National University of Singapore 学校（单位）

的高级研究学者□、访问学者□、博士后□ v、博士研究生□、硕士研究生□、

本科生□、大专生□、其他留学人员□

在我驻外使（领）馆报到日期 2016 年 08 月 15 日

注册入学日期 2015 年 09 月 15 日

毕（结）业日期 2016 年 09 月 14 日

拟回国日期 2016 年 09 月 20 日

毕（结）业证书名称 其它 号码

备注（留学经历描述）

留学回国人员签字：

经办人签字：

负责人签字：

教育（文化）处（组）公章

2016 年 09 月 06 日

第一联：交留学回国人员 教育部国际合作与交流司 2012 年制表

注意事项

1、本证明只为学成回国工作的留学人员开具。
2、本证明由我驻外使（领）馆教育（文化）处（组）在留学人员回国时填写，不得涂改。
3、本证明经使（领）馆教育（文化）处（组）经办人、负责人签字并在第一、第二联加盖公章方为有效。
4、第一联由留学人员保存，其他单位可查验原件，收存复印件，不得收取原件。

CONFIDENTIAL

Attn : To Whom It May Concern

DR WEN GUOXING

This is to certify that Dr Wen Guoxing is a Research Fellow of this University's Department of Electrical & Computer Engineering, Faculty of Engineering, having first joined the University on 15 September 2015. He is on a contract of service which will expire on 14 September 2016.

He is presently in receipt of a gross salary of S$ 4,500.00 per month.

DATED THIS THE FIFTEENTH DAY OF AUGUST 2016

OFFICE OF HUMAN RESOURCES ★ NUS National University of Singapore ★ OFFICE OF HUMAN RESOURCES

Tan Yong Suan (Ms)
for Vice President (Human Resources)
Tel: (65) 651 62331
E-mail: ohrsharedservices@nus.edu.sg

澳門特別行政區政府
Governo da Região Administrativa Especial de Macau
社會文化司司長辦公室
Gabinete do Secretário para os Assuntos Sociais e Cultura

批　示

本人同意澳門大學聘任文國興（中華人民共和國往來港澳通行證編號：C20058764）為科技學院兼職研究人員，合約期由2018年6月26日起至2018年9月25日止。

為辦理來澳門特別行政區工作的各項手續的需要，特立此批示，以茲證明。

二零一八年四月十六日。

澳門特別行政區政府
社會文化司司長

譚俊榮

澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU

15/03/2018

Dr. Wen Guoxing *(PASSPORT no. C20058764)*
Department of Electrical and Computer Engineering
Faculty of Science and Technology
University of Macau

Dear Dr. Wen,

On behalf of the University of Macau, I have the pleasure to confirm your appointment as **Post-doctoral Fellow** for the **FDCT Funded** project "Design of Discriminative Fuzzy Restricted Boltzmann Machines and their Applications" with the Faculty of Science and Technology according to article 7 item 4[1] and item 5[2] of the Personnel Statute of the University of Macau. The terms and conditions of this contract are set out hereunder:

1. Objective

Both parties agree to enter into this contract that serves for the FDCT Funded project "Design of Discriminative Fuzzy Restricted Boltzmann Machines and their Applications".

2. Duration

Both parties agree and consider this contract for a period of 3 months, beginning 26/06/2018 and ending 25/09/2018.

This contract shall ONLY take effect upon the fulfillment that you are a permanent resident of the Macao Special Administrative Region (SAR) or have obtained the approval and a valid identification document issued by relevant authorities of Macao for staying and working in the Macao SAR.

文国兴

Non-regular full-time

[1] The terms and conditions of this employment shall be stated in the contracts.

[2] The employees shall only be entitled to the rights stipulated in the contracts and subject to the obligations therein.

JA/RC/CI/alw

中國澳門氹仔 大學大馬路 電話：8822 8833 傳真：8822 8822
Avenida da Universidade, Taipa, Macau, China Tel: 8822 8833 Fax: 8822 8822

3. Duties

You shall engage in research task of the FDCT Funded project "Design of Discriminative Fuzzy Restricted Boltzmann Machines and their Applications" at the University of Macau, as assigned by and under the instructions and directions of the Head of the concerned unit.

You are exempted from the normal working time schedule for performing duties for the FDCT Funded project "Design of Discriminative Fuzzy Restricted Boltzmann Machines and their Applications", however your working time schedule shall be agreed by the concerned unit.

4. Remuneration

The taxable global remuneration of MOP71,499.00, including a basic accommodation allowance complied with the Chief Executive's Dispatch No. 88/2010, is payable by 3 equal monthly installments.

5. Special Rights

You are also entitled to the accessory rights as shown in the attached annex.

6. Outside Practice

You shall not engage in paid practice outside the University of Macau, unless you get a prior written approval from the Rector or his/her delegate(s). Please note that it is illegal for work permit holders (non-Macao residents) to engage in outside practice in Macao SAR.

文国兴

Non-regular full-time

JA/RC/CI/alw

7. Termination of Appointment

Upon the expiry of this contract, the appointment shall automatically terminate without any prior notice. For unilateral early termination of this contract, either party is required to give not less than one month's notice or indemnify the other Party for an equivalent payment corresponding to the unfulfilled notice days.

For and on behalf of
University of Macau

Yuing Guing Anchi Silva Aguiar
Head of Human Resources Section

I accept and confirm the above

文国兴 Wen Guoxing

Wen Guoxing

c.c.: Research & Development Administration Office

Non-regular full-time

JA/RC/CI/alw

澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU

ANNEX
REGARDING THE ACCESSORY RIGHTS TO THE CONTRACT BETWEEN THE UNIVERSITY OF MACAU AND DR. WEN GUOXING

Dr. Wen is entitled to the following accessory rights:

1. Medical benefits

The major medical benefits listed below are in accordance with the guidelines for medical benefits as stipulated in Appendix 6.3 of Chapter 6 of the Rules of the Personnel Affairs of the University of Macau:

- To be entitled to the medical benefits, 0.5% of the monthly remuneration during the period as stated in the existing contract will be deducted.
- The maximum total amount reimbursable for the medical expenses within the existing contract period shall be 20% of the total remuneration of the current contract less the hospitalization insurance premium (if applicable).
- For the medical expenses to be reimbursable, you shall bear 10% of the medical expenses incurred from each of the medical consultation and treatment, except those incurred from the medical and health care service provider on campus.
- Family members can also enjoy the medical benefits of both out-patient and hospitalization treatments if they can fulfill the terms and conditions specified in the above-mentioned Appendix 6.3.

2. 5 working days of paid leave which shall be taken within the period as stated in the existing contract. There shall be no deferment of and payment in lieu of unused paid leave.

3. Transportation reimbursement at a maximum of MOP3,500.00 for the returning trip to Shandong, China and it is payable upon the cessation of employment relation with the University. The reimbursement amount will be paid in accordance with actual expenditure, upon presentation of relevant supporting documents such as invoice and jet foil ticket.

University of Macau, 15/03/2018

For and on behalf of
University of Macau

Yuing Guing Anchi Silva Aguiar
Head of Human Resources Section

文国兴 Wen Guoxing
Wen Guoxing

Non-regular full-time

JA/RC/CI/alw

证书编号：B2UJYJH20170210

中青年拔尖人才
培育支持计划入选证书

文国光 同志

入选第 二 层次中青年拔尖人才

培育支持计划（聚英计划）

滨州学院

2018 年 3 月 20 日

证

为表彰滨州市自然科学优秀学

成果名称：基于神经网络的二阶非线性
的自适应领导–跟随者一致控

获奖作者：文国兴 陈俊龙 刘艳

中 共
滨州市委组织部

滨 州 市
人力资源和社会保障局

IEEE

IEEE Computational Intelligence Society

IEEE Transactions on Neural Networks and Learning Systems Outstanding Paper Award for 2014 (bestowed in 2017)

is presented to

C. L. P. Chen

For the paper co-authored with G. X. Wen, Y. J. Liu, and F. Y. Wang, entitled "Adaptive Consensus Control for a Class of Nonlinear Multiagent Time-delay Systems using Neural Networks," IEEE Transactions on Neural Networks and Learning Systems, Vol. 25, No. 6, pp. 1217-1226, 2014.

Pablo A. Estevez

Pablo A. Estevez
President
IEEE Computational Intelligence Society

FOUNDED NEW YORK 1884

July 2018

澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU

畢業證書

學生文國興修業期滿，考試及格，照章授予哲學博士學位（軟件工程）。

此證

公元二零一四年十一月十九日

CARTA DE CURSO

Certifica-se que
WEN GUOXING
concluiu com aproveitamento o curso
tendo-lhe sido conferido o grau de
DOUTOR em ENGENHARIA INFORMÁTICA
Macau, aos 19 de Novembro de 2014

CERTIFICATE

This is to certify that
having passed the examinations and having fulfilled all prescribed requirements
WEN GUOXING
has been awarded the degree of
DOCTOR OF PHILOSOPHY in SOFTWARE ENGINEERING
Macao, 19 November 2014

校 長
O Reitor
Rector

教務長
O Coordenador do Gabinete
de Assuntos Académicos
Registrar

14-30425

教育部留学服务中心

香港、澳门特别行政区
学历学位认证书

教留服认澳门[2015]00077号

文国兴，男，中国国籍，1977年2月15日生于山东省。

文国兴2011年9月起在澳门大学(Universidade de Macau)从事软件工程专业研究，论文通过，于2014年11月获得澳门大学颁发的毕业证书，并被授予哲学博士学位。

经核查，澳门大学系中国澳门特别行政区正规高等学校。

澳门特别行政区高等教育实行单证书制度。学生所获不同层次学位表明其具有相应的学历。

教育部留学服务中心

港澳台地区学历学位认证办公室

二〇一五年一月十六日

查询网址：www.cscse.edu.cn

教育部留学服务中心

香港、澳门特别行政区
学历学位认证书

教留服认澳门[2015]00077号

文国兴，男，中国国籍，1977年2月15日生于山东省。

文国兴2011年9月起在澳门大学(Universidade de Macau)从事软件工程专业研究，论文通过，于2014年11月获得澳门大学颁发的毕业证书，并被授予哲学博士学位。

经核查，澳门大学系中国澳门特别行政区正规高等学校。

澳门特别行政区高等教育实行单证书制度。学生所获不同层次学位表明其具有相应的学历。

教育部留学服务中心

港澳台地区学历学位认证办公室

二〇一五年一月十六日

查询网址：www.cscse.edu.cn

附件 2

推荐山东省有突出贡献的中青年专家“六公开”监督卡

推荐人选所在单位(盖章)：滨州学院　　　　2019 年 7 月 2 日

推荐人选姓名：文国兴

专业技术人员总数	1090	实际申报人数	5	推荐人选数量	2
“六公开”内　容	1. 公开推荐选拔范围 2. 公开推荐选拔条件 3. 公开推荐选拔数量	4. 公开推荐选拔程序 5. 公开推荐人选名单及排序 6. 公开推荐人选的评审材料			
如果认为单位做到了上述要求，且同意推荐人选申报，请在下面栏目中签名					
专业技术人员代表					
高丽	孙建武	[illegible]	范丽娟	[illegible]	
刘宇华	[illegible]	张艳梅	[illegible]	黄利国	
[illegible]	[illegible]	魏媛	[illegible]	[illegible]	
[illegible]	伊厚会	王学玲	[illegible]	[illegible]	
未签名人员及原因					
单位人事部门负责人	[illegible]				
单位领导	[illegible]				

备注：1. 专业技术人员总数少于 15 名的，由全体专业技术人员签名，未签名人员须列出名单并填写原因；超过（含）15 名的，须推选出不少于 15 名的代表签名。

2. 如违反“六公开”，请向本级、本部门（单位）人事部门反映，或邮寄至：济南市解放东路 16 号省人力资源社会保障厅人才开发处，邮编 250014。

山东省人力资源和社会保障厅

六公开卡.jpg

滨州学院人事处

证　　明

兹证明文国兴，男，身份证号：37230119770215075X，为我校理学院在职教师，为正式在编人员，特此证明。

滨州学院人事处

2019 年 7 月 03 日

教师证明.jpg

姓 名 文国兴

性 别 男 民 族 汉

出 生 1977 年 2 月 15 日

住 址 山东省滨州市滨城区渤海三路508号3号楼2单元102室

公民身份号码 37230119770215075X

中华人民共和国

居民身份证

签发机关 滨州市公安局滨城分局

有效期限 2006.04.19-2026.04.19

曲阜师范大学文件

曲师大校字〔2018〕75 号

曲阜师范大学
关于公布 2018 年研究生导师考核和增选结果的
通　　知

各学院（部），各部门,各单位:

根据《博士生指导教师资格审定和聘任工作实施细则》《硕士研究生指导教师资格审定和工作条例》《专业学位硕士研究生指导教师资格审定和工作条例》等相关规定，学校于 2018 年 7 月组织开展了研究生导师增选和考核工作。经各学院学位评定分委员会初审，学校组织专家复审，校学位评定委员会议决，9 名博士生导师、98 名学术型硕士生导师、96 名专业型硕士生导师

考核结果为合格；增选博士生导师 34 名、学术学位硕士生导师 157 名、专业学位硕士生导师 132 名。现将结果予以公布（见附件）。

如同一导师的考核合格专业和增选专业不同，以增选的专业为准。研究生导师聘期为 2018 年 1 月 1 日至 2021 年 12 月 31 日，若届中达到退休年龄的，聘期至退休之日。

附件：1. 2018 年博士生导师考核结果
2. 2018 年学术学位硕士生导师考核结果
3. 2018 年专业学位硕士生导师考核结果
4. 2018 年博士生导师增选名单
5. 2018 年学术学位硕士生导师增选名单
6. 2018 年专业学位硕士生导师增选名单

曲阜师范大学

2018 年 10 月 30 日

68	曹新华	工程硕士	电气工程	兼职
69	常伟	工程硕士	电气工程	兼职
70	丁方锋	工程硕士	电气工程	兼职
71	甘亚光	工程硕士	电气工程	兼职
72	郭然甲	工程硕士	电气工程	兼职
73	来庆亮	工程硕士	电气工程	兼职
74	李刚	工程硕士	电气工程	兼职
75	栗桂娜	工程硕士	电气工程	兼职
76	刘玲	工程硕士	电气工程	兼职
77	苗珍	工程硕士	电气工程	兼职
78	牟建	工程硕士	电气工程	兼职
79	彭飞	工程硕士	电气工程	兼职
80	邱芳	工程硕士	电气工程	兼职
81	邱建龙	工程硕士	电气工程	兼职
82	邵杰	工程硕士	电气工程	
83	文国兴	工程硕士	电气工程	兼职
84	闫绍敏	工程硕士	电气工程	
85	张安彩	工程硕士	电气工程	兼职
86	张成新	工程硕士	电气工程	
87	张可程	工程硕士	电气工程	兼职
88	崔新春	工程硕士	计算机技术	
89	侯林林	工程硕士	计算机技术	
90	马跃峰	工程硕士	计算机技术	
91	孙玉红	工程硕士	计算机技术	

白底照片.jpg

山东省自然科学基金

资助项目立项任务书

项目基本信息	项目名称	多智能体编队的优化控制				
	立项编号	ZR2018MF015			项目类别	面上项目
	执行期限	2018-03至2021-06			资助经费	14.00万元
	学科分类	最优控制			学科代码	F030113
项目承担人信息	姓名	文国兴	性别	男	身份证号	37230119770215075X
	电子邮箱	gxwen@live.cn			联系电话	15266765110
	单位名称	滨州学院			专业技术职务	讲师
	所在单位(院系)	数学系			主管部门	省教育厅
	所在省级以上重点实验室	无				

项目组成员（与申请书一致，不包含主持人）

姓名	职称	工作单位	任务分工	每年工作时间（月）	签名
王少英	讲师	滨州学院	理论研究	9	王少英
由红连	副教授	滨州学院	数学建模与理论推导	9	由红连
于海芳	副教授	滨州学院	算法验证与应用	9	于海芳
冯君	副教授	滨州学院	计算机仿真	9	冯君
麻连刚	讲师	滨州学院	数学建模与理论推导	9	麻连刚
高发亮	讲师	滨州学院	算法验证与应用	9	高发亮

需呈交科技报告（篇）

年度进展报告	最终(技术)报告(必须填，一般为1)
3	1

注：严格按照科技报告的有关规定呈交科技报告。项目执行中，年度或中期审核前应呈交进展报告；项目完成后三个月内、开展验收前，须呈交最终（技术）报告。未完成科技报告任务的，项目不予结题。

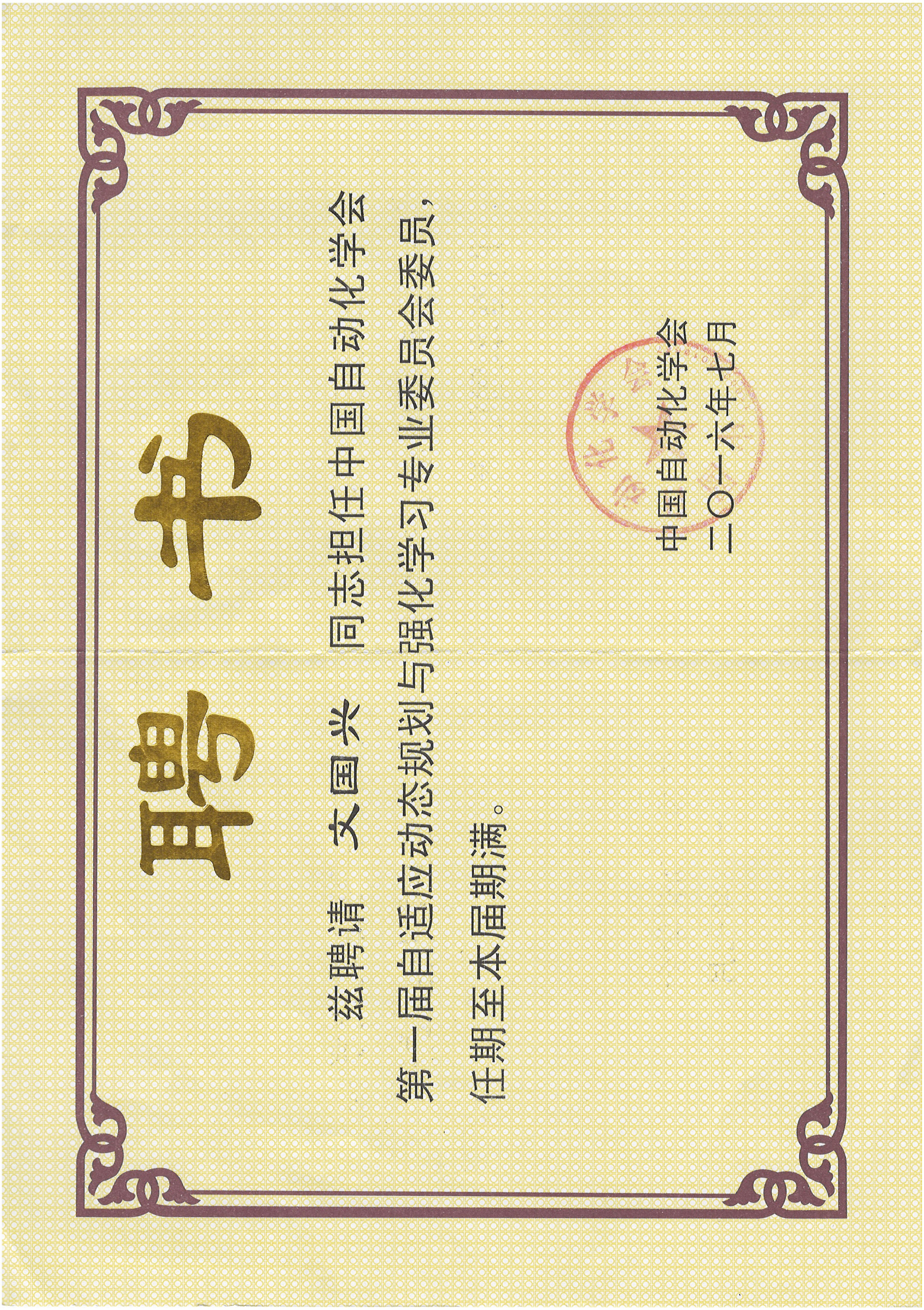
资助经费预算表（单位：万元）		
科目	预算经费	备注(计算依据与说明)
项目资助总额	14.00	
一、项目直接费用	12.00	
1、设备费	2.00	
(1)设备购置费	2.00	购买高性能户外电脑、无人机、打印机等实验设备。
(2)设备试制费	0.00	
(3)设备改造与租赁费	0.00	
2、材料费	0.00	
3、测试化验加工费	0.00	
4、燃料动力费	0.00	
5、差旅/会议/国际合作与交流费	4.00	参加国内外控制会议，如：中国CCC控制大会、台湾CAC国际控制会议所产生的差旅费及市内交通费。
6、出版/文献/信息传播/知识产权事务费	1.50	在国内、外期刊发表论文所产生费用，如：IEEE trans期刊，超过规定页数，每页收费160-200美元不等。
7、劳务费	1.50	用于外请的专家报告的劳务费及相关费用。
8、专家咨询费	2.00	外籍专家和国内专家咨询费用。每人次 800-200不等
9、其他支出	1.00	
二、项目间接经费 (比例:20%)	2.00	
1、绩效支出	1.30	
2、管理费	0.70	
3、房屋占用/日常水电气暖消耗	0.00	
三、自筹资金	0.00	

项目负责人承诺：本人接受山东省自然科学基金的资助，并将严格遵守山东省自然科学基金委员会关于资助项目管理和财务管理的各项规定，认真开展研究工作，按照项目申请书中的内容完成各项指标。按时报送有关材料，及时报告重大变动情况，对资助项目发表的论著和取得的研究成果按规定进行标注。

项目负责人签字：刘国兴　　　　2018年4月8日

依托单位审核意见	山东省自然科学基金委员会办公室审查意见
我单位同意承担该项目，将保证项目负责人及其研究队伍的稳定和项目实施所需的条件，严格遵守山东省自然科学基金委员会有关资助项目管理、财务等各项规定，并做好督促协调工作。 依托单位（公章） 2018年4月8日	（公章） 2018年4月9日

（正反面打印，一式三份）　　　　山东省自然科学基金委员会办公室2017年制



聘书

兹聘请　文国兴　同志担任中国自动化学会第一届自适应动态规划与强化学习专业委员会委员，任期至本届期满。

中国自动化学会

二〇一六年七月

中国自动化学会委员.jpg

滨州学院文件

滨院政聘字〔2019〕8号

关于聘任马国利等同志为校聘教授副教授职务的通知

各二级学院、部门，校直各单位：

兹聘任：

一、校聘教授

马国利　亓佩成　文国兴　刘金涛　刘　涛　孙景宽
杨红军　郑晶静　房吉敦　柳　明　姚　健　徐新生
韩春艳

二、校聘副教授

邢雪阳　张玉苗　张再旺　陈永彬　苟建霞　赵自国
赵英渊　赵晓光　郝　伟　郭瑞超　章夫正　谢振伟
穆文英　魏守才　魏德宸

以上同志聘任时间自2019年6月20日至2022年6月19日。

滨州学院

2019年6月21日

教授聘任文件.jpg.jpg

聘书

兹聘任 文国兴 为校聘 教授 ，聘期

自 2019 年 6 月 20日至 2022 年 6 月 19日。

滨州学院

校长：李长海

证书编号：BZUXP1003

2019 年 6 月20日

教授聘书.jpg.jpg