This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Adaptive Multigradient Recursive Reinforcement Learning Event-Triggered Tracking Control for Multiagent Systems

Hongyi Li, *Senior Member, IEEE*, Ying Wu, Mou Chen, *Senior Member, IEEE*, and Renquan Lu, *Senior Member, IEEE*

***Abstract*— This article proposes a fault-tolerant adaptive multigradient recursive reinforcement learning (RL) event-triggered tracking control scheme for strict-feedback discrete-time multiagent systems. The multigradient recursive RL algorithm is used to avoid the local optimal problem that may exist in the gradient descent scheme. Different from the existing event-triggered control results, a new lemma about the relative threshold event-triggered control strategy is proposed to handle the compensation error, which can improve the utilization of communication resources and weaken the negative impact on tracking accuracy and closed-loop system stability. To overcome the difficulty caused by sensor fault, a distributed control method is introduced by adopting the adaptive compensation technique, which can effectively decrease the number of online estimation parameters. Furthermore, by using the multigradient recursive RL algorithm with less learning parameters, the online estimation time can be effectively reduced. The stability of closed-loop multiagent systems is proved by using the Lyapunov stability theorem, and it is verified that all signals are semiglobally uniformly ultimately bounded. Finally, two simulation examples are given to show the availability of the presented control scheme.**

***Index Terms*— Event-triggered control, fault-tolerant control, multiagent systems, multigradient recursive reinforcement learning (RL) algorithm.**

I. Introduction

RECENTLY, the cooperative control problem of multiagent systems (MASs) has received considerable attention due to its widespread applications [1]–[7]. For example, an adaptive distributed fault-tolerant control approach with output regulation was proposed for linear MASs in [8]. Liu *et al.* [9] proposed a new distributed cooperative compound tracking scheme for the vehicular platoon using neural networks (NNs). The finite-time fault-tolerant distributed adaptive tracking control problem was solved for high-order nonlinear MASs in [10]. Wu *et al.* [11] proposed a cooperative adaptive control scheme by using the backstepping technique for high-order stochastic nonlinear MASs subject to dead zone. It is noted that the results in [8]–[11] mainly solved the weight vector problem. However, the ideal weight vector problem has not been considered in the above achievements. Therefore, it is important to design an adaptive control algorithm to search for the optical solution.

The gradient descent method, which can search the optimal solution, has received considerable attention [12]–[29]. For instance, Bai *et al.* [30] proposed an NN reinforcement learning (RL) control scheme for discrete-time nonstrict-feedback systems. A distributed fault-tolerant adaptive control scheme was put forward for MASs via gradient descent method in [31]. Yang and Jagannathan [32] proposed an adaptive critic control strategy for discrete-time systems via the RL algorithm. However, the methods in [12]–[32] searched the optimal solution from a single point, which is easy to cause the local optimal problem. To resolve it, Bai *et al.* [33] designed a novel multigradient recursive RL algorithm. For a family of nonlinear systems with input saturation, an adaptive multigradient recursive RL NN tracking control scheme was presented in [34], whereas the studied control methods in [33] and [34] need to adjust a large number of parameters online so that much estimation time would be wasted. Thus, we need to design an adaptive multigradient recursive RL algorithm that requires less online estimation time.

On the other hand, the inevitable limitation of network resources is considered to be the main cause of unexpected network-induced phenomenon. Therefore, the event-triggered mechanism as one of the effective ways to reduce the communication burden has attracted much interest [35]–[38]. For instance, Mu *et al.* [39] presented an event-triggered adaptive robust critic control strategy for a kind of nonlinear continuous-time systems. An event-triggered decentralized fuzzy adaptive control scheme was proposed for nonlinear large-scale systems in [40]. For a class of stochastic nonlinear MASs with unmeasurable states, Liang *et al.* [41] designed

Manuscript received November 25, 2020; revised April 9, 2021; accepted June 6, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 62033003, in part by the Local Innovative and Research Teams Project of Guangdong Special Support Program under Grant 2019BT02X353, and in part by the Innovative Research Team Program of Guangdong Province Science Foundation under Grant 2018B030312006. *(Corresponding author: Hongyi Li.)*

Hongyi Li and Renquan Lu are with the School of Automation, Guangdong Province Key Laboratory of Intelligent Decision and Cooperative Control, Guangdong University of Technology, Guangzhou 510006, China (e-mail: lihongyi2009@gmail.com; rqlu@gdut.edu.cn).

Ying Wu is with the College of Control Science and Engineering, Bohai University, Jinzhou 121013, China (e-mail: wuying_2018@163.com).

Mou Chen is with the College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China (e-mail: chenmou@nuaa.edu.cn).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TNNLS.2021.3090570.

Digital Object Identifier 10.1109/TNNLS.2021.3090570

an event-triggered fuzzy bipartite tracking control scheme, and the difficulty caused by unmeasurable states was solved by using the distributed reduced-order observer. The results in [39]–[41] considered the tradeoff between resource saving and control performance. It is known that an unnecessarily large event-triggered threshold will reduce the tracking accuracy and endanger the stability of the closed-loop systems. Therefore, Wang *et al.* [42] found an effective compensation method to weaken the impact of event-triggered threshold on the closed-loop systems' performance. Unfortunately, the measurement error in [42] needs to be bounded by a given value regardless of the size of control signal, which may degrade the system accuracy. Thus, it is meaningful to design a new compensation mechanism under the relative threshold event-triggered control.

The above analysis motivates us to present a fault-tolerant adaptive multigradient recursive RL event-triggered tracking control strategy for nonlinear strict-feedback MASs. The corresponding challenging problems at this work are listed as follows.

1) Different from the gradient descent method [30]–[32], the multigradient recursive RL algorithm is presented in this article, which can avoid the local optimal issue. In addition, the proposed approach with less learning parameters (LLPs) can decrease online estimation time, which is hard to be achieved in [33] and [34].
2) Under the framework of relative threshold event-triggered strategy, a lemma (see Lemma 1) is proposed, which can effectively handle the compensation error caused by the operator.
3) A new relative threshold compensation strategy is proposed for event-triggered mechanism, which can improve the tracking accuracy and system stability while reducing the communication burden.

The rest of this article is organized as follows. The priori knowledge is given in Section II. An event-triggered-based multigradient recursive RL control strategy is presented in Section III. Section IV shows some results with respect to practical and numerical examples. Finally, conclusions are given in Section V.

Notation: $\mathbb{R}^m$ $(m = 1, 2, \ldots, n)$ represents the m-dimensional Euclidean space. $|\bullet|$ and $\|\bullet\|$ stand for the absolute value and the Euclidean norm of vector or matrix, respectively. ∇ and grad denote the gradient at some point. $\triangle$ is the differentiate operation.

II. Preliminaries

A. Graph Theory

The information exchange among agents is usually represented by a digraph. The digraph is expressed as $\mathcal{G} = (\mathcal{T}, \mathcal{H}, \mathcal{A})$ with $\mathcal{T} = (1, \ldots, N)$ denoting the nonempty node set, $\mathcal{H} \subseteq \mathcal{T} \times \mathcal{T}$ representing the edge set, and $\mathcal{A} = [a_{i,j}] \in \mathbb{R}^{N\times N}$ being the relevant adjacency matrix. Edge $(\mathcal{T}_j, \mathcal{T}_i) \in \mathcal{H}$ denotes an edge from i to j. In addition, if the information flows from node j to i, $a_{i,j} > 0$; otherwise, $a_{i,j} = 0$. $\Theta = \text{diag}\{\theta_1, \ldots, \theta_N\}$ stands for the diagonal matrix with $\theta_i = \sum_{j=1}^{N} a_{i,j}$. $\mathcal{C} = \Theta - \mathcal{A}$ is the Laplacian matrix. $\mathcal{V}_i = \{\mathcal{T}_j | (\mathcal{T}_j, \mathcal{T}_i) \in \mathcal{H}, i \neq j\}$ stands for the neighbors set of the agents.

The augmented graph is expressed as $\bar{\mathcal{G}} = (\bar{\mathcal{T}}, \bar{\mathcal{H}})$ with $\bar{\mathcal{T}} = (\mathbf{0}, 1, \ldots, N)$, $\bar{\mathcal{H}} = \bar{\mathcal{T}} \times \bar{\mathcal{T}}$, where $\mathbf{0}$ is the signal of leader. $\varsigma_{i,\mathbf{0}} > 0$ means that the leader $\mathbf{0}$ can send information to node i; otherwise, $\varsigma_{i,\mathbf{0}} = 0$.

Assumption 1 [41]: For the leader–follower MASs, the digraph $\mathcal{G}$ contains at least a spanning tree with node $\mathbf{0}$ as the root.

B. Problem Formulation

This article considers the discrete-time MASs as

$$\begin{cases} s_{i,1}(k+1) = \tau_{i,1}s_{i,2}(k) + g_{i,1}(\bar{s}_{i,1}(k)) \\ s_{i,m}(k+1) = \tau_{i,m}s_{i,m+1}(k) + g_{i,m}(\bar{s}_{i,m}(k)) \\ s_{i,n}(k+1) = \tau_{i,n}\nu_i(k) + g_{i,n}(\bar{s}_{i,n}(k)) + d_i(k) \\ y_i(k) = F(s_{i,1}(k)) \end{cases} \tag{1}$$

where $\bar{s}_{i,m}(k) = [s_{i,1}(k), s_{i,2}(k), \ldots, s_{i,m}(k)]^T \in \mathbb{R}^m$ $(2 \leq m \leq n-1)$ and $\bar{s}_{i,n}(k) = [s_{i,1}(k), s_{i,2}(k), \ldots, s_{i,n}(k)]^T \in \mathbb{R}^n$ represent the state vectors. $g_{i,l}(\bar{s}_{i,l}(k))$ stand for the unknown smooth nonlinear functions, $\tau_{i,l} > 0$ denote the known parameters, and $d_i(k)$ represent the unknown external disturbances with $l = 1, \ldots, n$ and $i = 1, \ldots, N$. $\nu_i(k) \in R$ and $y_i(k) \in R$ denote the input and output of the MASs, respectively. We define $y_d(k)$ as the leader's dynamic. In addition, $F(s_{i,1}(k)) = \check{\zeta}_i(k)s_{i,1}(k) + \imath_i(k)$ stands for the sensor fault with $\check{\zeta}_i(k)$ and $\imath_i(k)$ being the parameters of sensor fault.

Based on the analysis of sensor fault in [43] and [44], the sensor fault model is introduced in discrete-time MASs. The parameters of the sensor fault considered in this article satisfy $\underline{\check{\zeta}} \leq \check{\zeta}_i(k) \leq 1$ and $\underline{\imath} \leq \imath_i(k) \leq \bar{\imath}$ with $\underline{\check{\zeta}} > 0$, $\underline{\imath}$ and $\bar{\imath}$ being the lower and upper bounds of $\imath_i(k)$, respectively. According to the above analysis, the following four types of faults are considered in this article.

1) If $\check{\zeta}_i(k) = 1$ and $\imath_i(k)$ is a constant, the fault is called the bias fault.
2) If $\check{\zeta}_i(k) = 1$ and $|\imath_i(k)| = \varsigma k$ with $0 < \varsigma \ll 1$, the fault is called the drift fault.
3) If $\check{\zeta}_i(k) = 1$ and $|\imath_i(k)| < \bar{\imath}$ with $\triangle\imath_i(k) \to 0$, the fault is named loss of accuracy.
4) If $\underline{\check{\zeta}} \leq \check{\zeta}_i(k) < 1$ and $\imath_i(k) = 0$, the fault is called loss of effectiveness.

Thus, the sensor fault is organized as

$$y_i(k) = s_{i,1}(k) + f_i(k) \tag{2}$$

with $f_i(k) = (\check{\zeta}_i(k) - 1)s_{i,1}(k) + \imath_i(k)$.

On the other hand, inspired by [42], a new relative threshold strategy-based event-triggered scheme is introduced. Fig. 1 shows the block diagram of control scheme, and we define the transmission error as

$$\check{\upsilon}_i(k) = u_i^t(k) - u_i(k) = u_i(k_b) - u_i(k), \quad k_b \leq k < k_{b+1} \tag{3}$$

where $u_i(k)$ denotes the current feedback control signal and $u_i^t(k) = u_i(k_b)$ represents the last transmitted control signal

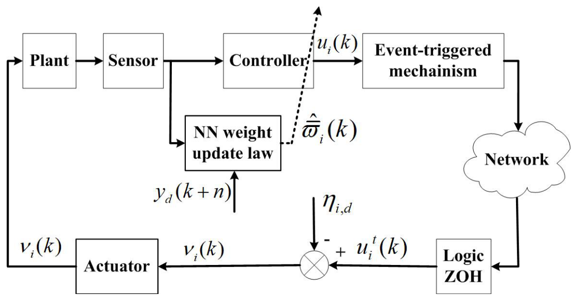


Fig. 1. Block diagram of the event-triggered adaptive control scheme.

with k_b standing for the triggering instant and $b \in N$. Then, the relative threshold control condition is designed as

$$|\breve{\upsilon}_i(k)| \geq \check{\eta}_i(|u_i(k)| + \eta_i) \tag{4}$$

where $0 < \check{\eta}_i < 1$ and $\eta_i > 0$ are designed parameters.

Based on (3) and (4), it yields

$$k_{b+1} = \min\{k \in N | k > k_b, \quad |\breve{\upsilon}_i(k)| \geq \check{\eta}_i(|u_i(k)| + \eta_i)\}.$$

In addition, to degrade the influence of event-triggered threshold on the performance of closed-loop MASs, a designed operator is introduced into the actuator as follows:

$$\nu_i(k) = u_i^t(k) - \eta_{i,d}$$

where $\eta_{i,d}$ will be designed later.

Assumption 2 [31], [45]: The leader $y_d(k)$ and its differences are assumed as bounded functions.

Assumption 3 [46]: The external disturbance $|d_i(k)| \leq \bar{d}$ is assumed to be bounded, and $\bar{d} > 0$ is a constant.

Lemma 1: For a positive constant $\breve{\sigma}_i$ and a time-varying parameter $|\varphi_i(k)| \leq 1$, one has

$$|\varphi_i(k)\check{\eta}_i\eta_i - \eta_{i,d}| < \check{\eta}_i\eta_i, \quad |\varphi_i(k)| < 1$$
$$|\varphi_i(k)\check{\eta}_i\eta_i - \eta_{i,d}| \leq 0.2785\breve{\sigma}_i, \quad |\varphi_i(k)| = 1$$

where $\eta_{i,d}$ is designed in (27).

Proof: Since $|\varphi_i(k)| \leq 1$, we discuss it in five cases as follows.

1) When $0 < \varphi_i(k) < 1$, we can get $\text{sgn}(\varphi_i(k)) = 1$, and thus, it follows that

$$|\varphi_i(k)\check{\eta}_i\eta_i - \eta_{i,d}| = \check{\eta}_i\eta_i\left|\varphi_i(k) - \tanh\left(\frac{\check{\eta}_i\eta_i}{\breve{\sigma}_i}\right)\right| \leq \check{\eta}_i\eta_i.$$

2) When $\varphi_i(k) = 0$, we can get $\text{sgn}(\varphi_i(k)) = 0$, and thus, it follows that

$$|\varphi_i(k)\check{\eta}_i\eta_i - \eta_{i,d}| = 0.$$

3) When $-1 < \varphi_i(k) < 0$, we can get $\text{sgn}(\varphi_i(k)) = -1$, and thus, it follows that

$$|\varphi_i(k)\check{\eta}_i\eta_i - \eta_{i,d}| = \check{\eta}_i\eta_i\left|\varphi_i(k) + \tanh\left(\frac{\check{\eta}_i\eta_i}{\breve{\sigma}_i}\right)\right| \leq \check{\eta}_i\eta_i.$$

4) When $\varphi_i(k) = 1$, we can get $\text{sgn}(\varphi_i(k)) = 1$, and thus, according to [36], it follows that

$$|\varphi_i(k)\check{\eta}_i\eta_i - \eta_{i,d}| = \left|\check{\eta}_i\eta_i - \check{\eta}_i\eta_i\tanh\left(\frac{\check{\eta}_i\eta_i}{\breve{\sigma}_i}\right)\right| \leq 0.2785\breve{\sigma}_i.$$

5) When $\varphi_i(k) = -1$, we can get $\text{sgn}(\varphi_i(k)) = -1$, and thus, one has

$$|\varphi_i(k)\check{\eta}_i\eta_i - \eta_{i,d}| = \left|-\check{\eta}_i\eta_i + \check{\eta}_i\eta_i\tanh\left(\frac{\check{\eta}_i\eta_i}{\breve{\sigma}_i}\right)\right| \leq 0.2785\breve{\sigma}_i.$$

Then, based on the above proof, we can get the conclusion of Lemma 1.

According to [47] and (2), one has

$$\begin{cases} s_{i,1}(k+n) = \tau_{i,1}s_{i,2}(k+n-1) + G_{i,1}(\bar{s}_{i,n}(k)) \\ s_{i,m}(\breve{k}-m) = \tau_{i,m}s_{i,m+1}(k+n-m) \\ \qquad\qquad + G_{i,m}(\bar{s}_{i,n}(k)) \\ s_{i,n}(k+1) = \tau_{i,n}\nu_i(k) + G_{i,n}(\bar{s}_{i,n}(k)) + d_i(k) \\ y_i(k) = s_{i,1}(k) + f_i(k) \end{cases} \tag{5}$$

where $\breve{k} = k+n+1$ and $G_{i,1}(\cdot)$, $G_{i,m}(\cdot)$, and $G_{i,n}(\cdot)$ are new functions with $2 \leq m \leq n-1$.

The control goal of this article is to develop an adaptive multigradient recursive RL event-triggered scheme to ensure that all signals of the closed-loop MASs are semiglobally uniformly ultimately bounded (SGUUB).

The distributed errors and the change of coordinates are defined as

$$\begin{cases} \xi_{i,1}(k+n) = \sum_{j\in\mathcal{V}_i} a_{i,j}(y_i(k+n) - y_j(k+n)) \\ \qquad + \varsigma_{i,\mathbf{0}}(y_i(k+n) - y_d(k+n)) \\ \xi_{i,m}(\breve{k}-m) = s_{i,m}(k+n+1-m) - \breve{\varrho}_{i,m-1}(k) \\ \xi_{i,n}(k+1) = s_{i,n}(k+1) - \breve{\varrho}_{i,n-1}(k) \end{cases} \tag{6}$$

with $\varsigma_{i,\mathbf{0}} \geq 0$ standing for the pinning gain and $\breve{\varrho}_{i,m-1}(k)$ and $\breve{\varrho}_{i,n-1}(k)$ being the virtual control signals.

C. Radial Basis Function NNs

The unknown nonlinear function $\varrho(\Lambda)$ is estimated by using the following radial basis function (RBF) NN:

$$\varrho(\Lambda) = \varpi^{*T}\phi(\Lambda)$$

with $\varpi^* \in \mathbb{R}^\beta$ standing for the weight vector and $\phi(\Lambda) = [\phi_1(\Lambda), \phi_2(\Lambda), \ldots, \phi_\beta(\Lambda)]^T$ denoting the basis function vector, where $\beta > 1$ represents the node number of NNs. We choose the Gaussian function form of $\phi_i(\Lambda)$ as

$$\phi_i(\Lambda) = \exp\left[-\frac{(\Lambda - \check{\nu}_i)^T(\Lambda - \check{\nu}_i)}{\check{\tau}_i^2}\right], \quad i = 1, 2, \ldots, \beta$$

where $\check{\nu}_i$ stands for the receptive field center and $\check{\tau}_i$ represents the Gaussian function width.

Obviously, we can get

$$\phi_i^T(\Lambda)\phi_i(\Lambda) \leq \beta.$$

Then, based on [48]–[53], for a constant $\bar{\varepsilon} > 0$ and any unknown function $\varrho(\Lambda)$ over a compact set, one has

$$\varrho(\Lambda) = \varpi^T\phi(\Lambda) + \varepsilon(\Lambda)$$

where ϖ represents the ideal weight vector and $\varepsilon(\Lambda)$ stands for the approximation error with $|\varepsilon(\Lambda)| \leq \bar{\varepsilon}$.

III. Adaptive Multigradient Recursive RL Control Design

A. Designs of Strategic Utility Function and Critic NNs

Based on [31], we define

$$\Upsilon_i(k) = \begin{cases} 0, & |\xi_{i,1}(k)| \le \breve{\upsilon} \\ 1, & |\xi_{i,1}(k)| > \breve{\upsilon} \end{cases}$$

where $\breve{\upsilon}$ denotes the tracking performance threshold and $\Upsilon_i(k)$ represents the tracking performance index.

Then, from [54], the long-term strategic utility function (SUF) can be designed as

$$Q_i(k) = \Upsilon_i(k+1) + \breve{\delta}_i \Upsilon_i(k+2) + \breve{\delta}_i^2 \Upsilon_i(k+3) + \dots$$

where $\breve{\delta}_i > 0$ denotes a designed weighting parameter.

Since $Q_i(k)$ is unknown, critic NNs are used to approximate it, and one has

$$Q_i(k) = \varpi_{i,o}^T \phi_{i,o}(k) + \varepsilon_{i,o}(k)$$

where $\varpi_{i,o}$ stands for the ideal weight vector and $\varepsilon_{i,o}(k)$ represents the approximation error.

Based on the idea of LLPs in [55], it yields

$$\hat{Q}_i(k) = \hat{\bar{\varpi}}_{i,o}(k)\bar{\phi}_{i,o}(k) \tag{7}$$

with $\hat{Q}_i(k)$ being the estimation of $Q_i(k)$, $\hat{\bar{\varpi}}_{i,o}$ being the estimation of $\bar{\varpi}_{i,o} = \|\varpi_{i,o}\|$, and $\bar{\phi}_{i,o}(k) = \|\phi_{i,o}(k)\|$.

Remark 1: To obtain better control performance, the number of NN nodes in the hidden layer may be large, and then, more computation time is required. Therefore, the method with LLPs that adjusts the Euclidean norm of weights is used in this article to reduce the online computation time.

Similar to [56]–[58], we define the prediction error as

$$\xi_{i,o}(k) = \breve{\delta}_i \hat{Q}_i(k) - [\hat{Q}_i(k-1) - \Upsilon_i(k)].$$

Then, we define the cost function (CF) as $\Theta_{i,o}(k) = \xi_{i,o}^2(k)/2$, and the gradient can be derived as

$$\begin{aligned}\nabla\hat{\bar{\varpi}}_{i,o}(k) &= \frac{\partial \Theta_{i,o}(k)}{\partial \hat{\bar{\varpi}}_{i,o}(k)} \\ &= \breve{\delta}_i \bar{\phi}_{i,o}(k)[\breve{\delta}_i \hat{Q}_i(k) - \hat{Q}_i(k-1) + \Upsilon_i(k)].\end{aligned}$$

According to [34], the multigradient can be obtained as

$$\begin{aligned}\mathrm{grad}(\rho, \Theta_{i,o}(k)) = \sum_{l=1}^{\rho} &\breve{\delta}_i \bar{\phi}_{i,o}(k-l+1) \\ &\times [\breve{\delta}_i \chi_{i,o}(k-l+1) - \chi_{i,o}(k-l) \\ &+ \Upsilon_i(k-l+1)]\end{aligned} \tag{8}$$

where ρ denotes the gradient length, $\chi_{i,o}(k-l+1) = \hat{\bar{\varpi}}_{i,o}(k)\bar{\phi}_{i,o}(k-l+1)$, and $\chi_{i,o}(k-l) = \hat{\bar{\varpi}}_{i,o}(k)\bar{\phi}_{i,o}(k-l)$.

The multigradient recursive updating law can be designed as

$$\begin{aligned}\hat{\bar{\varpi}}_{i,o}(k+1) = \hat{\bar{\varpi}}_{i,o}(k) - \zeta_{i,o} \sum_{l=1}^{\rho} &\breve{\delta}_i \bar{\phi}_{i,o}(k-l+1) \\ &\times [\breve{\delta}_i \chi_{i,o}(k-l+1) - \chi_{i,o}(k-l) \\ &+ \Upsilon_i(k-l+1)]\end{aligned} \tag{9}$$

where $\zeta_{i,o} > 0$ represents the convergence factor.

B. Design of Adaptive RL Controller

By using the direct adaptive control method, an event-triggered-based distributed NN multigradient recursive RL controller will be designed in this section.

Step 1: Based on (5) and (6), it yields

$$\begin{aligned}\xi_{i,1}(k+n) = &(\theta_i + \varsigma_{i,\mathbf{0}})(G_{i,1}(\bar{s}_{i,n}(k)) + \tau_{i,1}s_{i,2}(k+n-1) \\ &+ f_i(k+n)) - \sum_{j\in\mathcal{V}_i} a_{i,j} \\ &\times (G_{j,1}(\bar{s}_{j,n}(k)) + \tau_{j,1}s_{j,2}(k+n-1) \\ &+ f_j(k+n)) - \varsigma_{i,\mathbf{0}} y_d(k+n).\end{aligned} \tag{10}$$

Inspired by the idea of direct adaptive control method in [59], the desired control signal is designed as

$$\begin{aligned}\varrho_{i,1}(k) = -\frac{1}{\tau_{i,1}(\theta_i + \varsigma_{i,\mathbf{0}})}\Bigg[&(\theta_i + \varsigma_{i,\mathbf{0}})(G_{i,1}(\bar{s}_{i,n}(k) + f_i(k+n)) \\ &- \sum_{j\in\mathcal{V}_i} a_{i,j}(G_{j,1}(\bar{s}_{j,n}(k)) \\ &+ \tau_{j,1}s_{j,2}(k+n-1) + f_j(k+n)) \\ &- \varsigma_{i,\mathbf{0}} y_d(k+n)\Bigg].\end{aligned}$$

Since the desired control signal contains the unknown continuous functions $G_{i,1}(\bar{s}_{i,n}(k))$, $G_{j,1}(\bar{s}_{j,n}(k))$, $f_i(k+n)$, and $f_j(k+n)$, the RBF NNs are employed to approximate it, that is,

$$\varrho_{i,1}(k) = \varpi_{i,1}^T \phi_{i,1}(\omega_{i,1}(k)) + \varepsilon_{i,1}(k) \tag{11}$$

where $\omega_{i,1}(k) = [\bar{s}_{i,n}^T(k), y_d(k+n)]^T$. Then, based on (6), (10), and (11), one can obtain

$$\begin{aligned}\xi_{i,1}(k+n) = (\theta_i + \varsigma_{i,\mathbf{0}})\tau_{i,1}(&\xi_{i,2}(k+n-1) \\ &+ \breve{\varrho}_{i,1}(k) - \varrho_{i,1}(k)).\end{aligned}$$

Remark 2: The existing results use dynamic programming methods to solve nonlinear optimization problems, and however, dynamic programming methods have two disadvantages: OFF-line computation and large computation. Therefore, the RL algorithm is introduced in this article and the NNs are used to approximate nonlinear functions online. This algorithm can make up for the shortcomings of dynamic programming.

By using the approach with LLPs, we design the virtual control signal as

$$\breve{\varrho}_{i,1}(k) = \hat{\bar{\varpi}}_{i,1}(k)\bar{\phi}_{i,1}(\omega_{i,1}(k)) \tag{12}$$

with $\hat{\bar{\varpi}}_{i,1}$ being the estimate of $\bar{\varpi}_{i,1} = \|\varpi_{i,1}\|$ and $\bar{\phi}_{i,1} = [\|\phi_{1,1}\|, \dots, \|\phi_{N,1}\|]^T$.

In the light of (11) and (12), we have

$$\begin{aligned}\xi_{i,1}(k+n) = &(\theta_i + \varsigma_{i,\mathbf{0}})\tau_{i,1}\xi_{i,2}(k+n-1) \\ &+ (\theta_i + \varsigma_{i,\mathbf{0}})\tau_{i,1}\tilde{\bar{\varpi}}_{i,1}(k)\bar{\phi}_{i,1}(\omega_{i,1}(k)) \\ &+ (\theta_i + \varsigma_{i,\mathbf{0}})\tau_{i,1}\epsilon_{i,1}(k)\end{aligned} \tag{13}$$

where $\tilde{\hat{\bar{\varpi}}}_{i,1}(k) = \hat{\bar{\varpi}}_{i,1}(k) - \bar{\varpi}_{i,1}$ and $\epsilon_{i,1}(k) = \bar{\varpi}_{i,1}\bar{\phi}_{i,1}(\omega_{i,1}(k)) - \varpi_{i,1}^T\phi_{i,1}(\omega_{i,1}(k)) - \varepsilon_{i,1}(k)$. Then, (13) can be converted to

$$\begin{aligned}\xi_{i,1}(k+1) = {} & (\theta_i + \varsigma_{i,\mathbf{0}})\tau_{i,1}\xi_{i,2}(k)\\ & + (\theta_i + \varsigma_{i,\mathbf{0}})\tau_{i,1}\tilde{\hat{\bar{\varpi}}}_{i,1}(k_1)\bar{\phi}_{i,1}(\omega_{i,1}(k_1))\\ & + (\theta_i + \varsigma_{i,\mathbf{0}})\tau_{i,1}\epsilon_{i,1}(k_1)\end{aligned} \tag{14}$$

where $k_1 = k - n + 1$. In addition, we define

$$\xi_{i,c1}(k) = \hat{\bar{\varpi}}_{i,1}(k_1)\bar{\phi}_{i,1}(\omega_{i,1}(k_1)) + [\hat{Q}_i(k) - Q_{i,d}(k)]$$

with $Q_{i,d}(k)$ denoting the desired SUF and being usually chosen as "0" from [34]. Thus, we design the CF as

$$\Theta_{i,1}(k) = \frac{\xi_{i,c1}^2(k)}{2}.$$

Then, the gradient is expressed as

$$\begin{aligned}\nabla\hat{\bar{\varpi}}_{i,1}(k_1) &= \frac{\partial\Theta_{i,1}(k)}{\partial\hat{\bar{\varpi}}_{i,1}(k_1)}\\ &= \bar{\phi}_{i,1}(\omega_{i,1}(k_1))[\hat{\bar{\varpi}}_{i,1}(k_1)\bar{\phi}_{i,1}(\omega_{i,1}(k_1)) + \hat{Q}_i(k)].\end{aligned}$$

Similar to (8), one can obtain

$$\begin{aligned}\mathrm{grad}(\rho, \Theta_{i,1}(k)) = {} & \sum_{l=1}^{\rho}\bar{\phi}_{i,1}(\omega_{i,1}(k_1 - l + 1))\\ & \times [\chi_{i,1}(k_1 - l + 1) + \chi_{i,o}(k - l + 1)]\end{aligned} \tag{15}$$

where $\chi_{i,1}(k_1 - l + 1) = \hat{\bar{\varpi}}_{i,1}(k_1)\bar{\phi}_{i,1}(\omega_{i,1}(k_1 - l + 1))$. Then, the multigradient recursive updating law can be designed as

$$\begin{aligned}\hat{\bar{\varpi}}_{i,1}(k+1) = {} & \hat{\bar{\varpi}}_{i,1}(k_1) - \zeta_{i,1}\sum_{l=1}^{\rho}\bar{\phi}_{i,1}(\omega_{i,1}(k_1 - l + 1))\\ & \times [\chi_{i,1}(k_1 - l + 1) + \chi_{i,o}(k - l + 1)]\end{aligned} \tag{16}$$

where $\zeta_{i,1} > 0$ is the convergence factor.

Step m ($2 \le m \le n-1$): In the light of (5) and (6), we have

$$\begin{aligned}\xi_{i,m}(k + n - m + 1) = {} & G_{i,m}(\bar{s}_{i,n}(k)) + \tau_{i,m}s_{i,m+1}\\ & \times (k + n - m) - \check{\varrho}_{i,m-1}(k).\end{aligned}$$

The desired control signal is designed as

$$\varrho_{i,m}(k) = -\frac{1}{\tau_{i,m}}(G_{i,m}(\bar{s}_{i,n}(k)) - \check{\varrho}_{i,m-1}(k)).$$

Similar to (11), one has

$$\varrho_{i,m}(k) = \varpi_{i,m}^T\phi_{i,m}(\omega_{i,m}(k)) + \varepsilon_{i,m}(k) \tag{17}$$

where $\omega_{i,m}(k) = [\bar{s}_{i,n}^T(k), \check{\varrho}_{i,m-1}(k)]^T$. Then, based on (6), one yields

$$\begin{aligned}\xi_{i,m}(k + n - m + 1) = {} & \tau_{i,m}(\xi_{i,m+1}(k + n - m)\\ & + \check{\varrho}_{i,m}(k) - \varrho_{i,m}(k)).\end{aligned} \tag{18}$$

Inspired by the method with LLPs, we design the virtual control signal $\check{\varrho}_{i,m}(k)$ as

$$\check{\varrho}_{i,m}(k) = \hat{\bar{\varpi}}_{i,m}(k)\bar{\phi}_{i,m}(\omega_{i,m}(k)) \tag{19}$$

where $\hat{\bar{\varpi}}_{i,m}$ is the estimation of $\bar{\varpi}_{i,m} = \|\varpi_{i,m}\|$ and $\bar{\phi}_{i,m} = [\|\phi_{1,m}\|, \ldots, \|\phi_{N,m}\|]^T$.

Substituting (17) and (19) into (18), it gives

$$\begin{aligned}\xi_{i,m}(k + n - m + 1) = {} & \tau_{i,m}\xi_{i,m+1}(k + n - m) + \tau_{i,m}\\ & \times \tilde{\hat{\bar{\varpi}}}_{i,m}(k)\bar{\phi}_{i,m}(\omega_{i,m}(k)) + \tau_{i,m}\epsilon_{i,m}(k),\end{aligned}$$

where $\tilde{\hat{\bar{\varpi}}}_{i,m}(k) = \hat{\bar{\varpi}}_{i,m}(k) - \bar{\varpi}_{i,m}$ and $\epsilon_{i,m}(k) = \bar{\varpi}_{i,m}\bar{\phi}_{i,m}(\omega_{i,m}(k)) - \varpi_{i,m}^T\phi_{i,m}(\omega_{i,m}(k)) - \varepsilon_{i,m}(k)$.

Similar to (14), we can obtain

$$\begin{aligned}\xi_{i,m}(k+1) = {} & \tau_{i,m}\xi_{i,m+1}(k) + \tau_{i,m}\\ & \times \tilde{\hat{\bar{\varpi}}}_{i,m}(k_m)\bar{\phi}_{i,m}(\omega_{i,m}(k_m)) + \tau_{i,m}\epsilon_{i,m}(k_m)\end{aligned}$$

where $k_m = k - n + m$. Then, we define

$$\xi_{i,cm}(k) = \hat{\bar{\varpi}}_{i,m}(k_m)\bar{\phi}_{i,m}(\omega_{i,m}(k_m)) + [\hat{Q}_i(k) - Q_{i,d}(k)].$$

Thus, we design the CF as

$$\Theta_{i,m}(k) = \frac{\xi_{i,cm}^2(k)}{2}.$$

The gradient can be derived as

$$\begin{aligned}\nabla\hat{\bar{\varpi}}_{i,m}(k_m) &= \frac{\partial\Theta_{i,m}(k)}{\partial\hat{\bar{\varpi}}_{i,m}(k_m)}\\ &= \bar{\phi}_{i,m}(\omega_{i,m}(k_m))\\ &\quad \times [\hat{\bar{\varpi}}_{i,m}(k_m) \times \bar{\phi}_{i,m}(\omega_{i,m}(k_m)) + \hat{Q}_i(k)].\end{aligned}$$

Similar to (15), it yields

$$\begin{aligned}\mathrm{grad}(\rho, \Theta_{i,m}(k)) = {} & \sum_{l=1}^{\rho}\bar{\phi}_{i,m}(\omega_{i,m}(k_m - l + 1))\\ & \times [\chi_{i,m}(k_m - l + 1) + \chi_{i,o}(k - l + 1)]\end{aligned} \tag{20}$$

where $\chi_{i,m}(k_m - l + 1) = \hat{\bar{\varpi}}_{i,m}(k)\bar{\phi}_{i,m}(\omega_{i,m}(k_m - l + 1))$. Then, the multigradient recursive updating law can be designed as

$$\begin{aligned}\hat{\bar{\varpi}}_{i,m}(k+1) = {} & \hat{\bar{\varpi}}_{i,m}(k_m) - \zeta_{i,m}\sum_{l=1}^{\rho}\bar{\phi}_{i,m}(\omega_{i,m}(k_m - l + 1))\\ & \times [\chi_{i,m}(k_m - l + 1) + \chi_{i,o}(k - l + 1)]\end{aligned} \tag{21}$$

where $\zeta_{i,m} > 0$ is the convergence factor.

Step n: In the light of (5) and (6), we have

$$\begin{aligned}\xi_{i,n}(k+1) = {} & G_{i,n}(\bar{s}_{i,n}(k)) + \tau_{i,n}\nu_{i,n}(k) + d_i(k)\\ & - \check{\varrho}_{i,n-1}(k).\end{aligned} \tag{22}$$

For the convenience of controller design, the controller $u_i(k)$ is used to express the input signal $\nu_i(k)$ of MASs. According to the event-triggered condition in (4), there is a time-varying parameter $\varphi_i(k)$ that satisfies $|\varphi_i(k)| \le 1$, which makes the following equation holds:

$$u_i^t(k) = (1 + \varphi_i(k)\check{\eta}_i)u_i(k) + \varphi_i(k)\check{\eta}_i\eta_i.$$

Then, $\nu_i(k)$ can be rewritten as

$$\nu_i(k) = (1 + \varphi_i(k)\check{\eta}_i)u_i(k) + \varphi_i(k)\check{\eta}_i\eta_i - \eta_{i,d}. \tag{23}$$

Substituting (23) into (22), it gives

$$\begin{aligned}\xi_{i,n}(k+1) = {} & G_{i,n}(\bar{s}_{i,n}(k)) + \tau_{i,n}((1 + \varphi_i(k)\check{\eta}_i)\\ & \times u_i(k) + \varphi_i(k)\check{\eta}_i\eta_i - \eta_{i,d})\\ & + d_i(k) - \check{\varrho}_{i,n-1}(k).\end{aligned}$$

The desired control signal is designed as

$$\varrho_{i,n}(k) = -\frac{1}{\tau_{i,n}(1+\varphi_i(k)\breve{\eta}_i)}(G_{i,n}(\bar{s}_{i,n}(k)) - \breve{\varrho}_{i,n-1}(k)).$$

Following the procedure in Steps 1 and m, one has

$$\varrho_{i,n}(k) = \varpi_{i,n}^T\phi_{i,n}(\omega_{i,n}(k)) + \varepsilon_{i,n}(k) \tag{24}$$

where $\omega_{i,n}(k) = [\bar{s}_{i,n}^T(k), \breve{\varrho}_{i,n-1}(k)]^T$. Then, we can obtain

$$\begin{aligned}\xi_{i,n}(k+1) &= \tau_{i,n}(1+\varphi_i(k)\breve{\eta}_i)(u_i(k) - \varrho_{i,n}(k)) \\ &\quad + \tau_{i,n}\varphi_i(k)\breve{\eta}_i\eta_i - \tau_{i,n}\eta_{i,d} + d_i(k).\end{aligned} \tag{25}$$

Inspired by the method with LLPs, we design the actual control signal as

$$u_i(k) = \hat{\bar{\varpi}}_{i,n}(k)\bar{\phi}_{i,n}(\omega_{i,n}(k)) \tag{26}$$

with $\hat{\bar{\varpi}}_{i,n}$ being the estimate of $\bar{\varpi}_{i,n} = \|\varpi_{i,n}\|$ and $\bar{\phi}_{i,n} = [\|\phi_{1,n}\|, \ldots, \|\phi_{N,n}\|]^T$. In addition, $\eta_{i,d}$ is designed as

$$\eta_{i,d} = \breve{\eta}_i\eta_i \operatorname{sgn}(\varphi_i(k)) \tanh\left(\frac{\breve{\eta}_i\eta_i}{\breve{\sigma}_i}\right) \tag{27}$$

where $\breve{\sigma}_i > 0$ is defined in Lemma 1 and $\operatorname{sgn}(0) = 0$.

Substituting (24), (26), and (27) into (25), one has

$$\begin{aligned}\xi_{i,n}(k+1) &= \tau_{i,n}(1+\varphi_i(k)\breve{\eta}_i)\tilde{\bar{\varpi}}_{i,n}(k)\bar{\phi}_{i,n}(\omega_{i,n}(k)) \\ &\quad + \tau_{i,n}\pi_i(k) + \tau_{i,n}\epsilon_{i,n} + d_i(k)\end{aligned}$$

where $\tilde{\bar{\varpi}}_{i,n}(k) = \hat{\bar{\varpi}}_{i,n}(k) - \bar{\varpi}_{i,n}$, $\pi_i(k) = \varphi_i(k)\breve{\eta}_i\eta_i - \eta_{i,d}$ and $\epsilon_{i,n}(k) = (1+\varphi_i(k)\breve{\eta}_i)\bar{\varpi}_{i,n}\bar{\phi}_{i,n}(\omega_{i,n}(k)) - (1+\varphi_i(k)\breve{\eta}_i)\varpi_{i,n}^T\phi_{i,n}(\omega_{i,n}(k)) - (1+\varphi_i(k)\breve{\eta}_i)\varepsilon_{i,n}(k)$.

Then, we define

$$\xi_{i,cn}(k) = \hat{\bar{\varpi}}_{i,n}(k)\bar{\phi}_{i,n}(\omega_{i,n}(k)) + [\hat{Q}_i(k) - Q_{i,d}(k)].$$

Thus, we design the CF as

$$\Theta_{i,n}(k) = \frac{\xi_{i,cn}^2(k)}{2}.$$

The gradient can be derived as

$$\begin{aligned}\nabla\hat{\bar{\varpi}}_{i,n}(k) &= \frac{\partial\Theta_{i,n}(k)}{\partial\hat{\bar{\varpi}}_{i,n}(k)} \\ &= \bar{\phi}_{i,n}(\omega_{i,n}(k))[\hat{\bar{\varpi}}_{i,n}(k)\bar{\phi}_{i,n}(\omega_{i,n}(k)) + \hat{Q}_i(k)].\end{aligned}$$

Similar to (20), it yields

$$\begin{aligned}\operatorname{grad}(\rho, \Theta_{i,n}(k)) &= \sum_{l=1}^{\rho}\bar{\phi}_{i,n}(\omega_{i,n}(k-l+1)) \\ &\quad \times [\chi_{i,n}(k-l+1) + \chi_{i,o}(k-l+1)]\end{aligned}$$

where $\chi_{i,n}(k-l+1) = \hat{\bar{\varpi}}_{i,n}(k)\bar{\phi}_{i,n}(\omega_{i,n}(k-l+1))$. Then, the multigradient recursive updating law can be designed as

$$\begin{aligned}\hat{\bar{\varpi}}_{i,n}(k+1) &= \hat{\bar{\varpi}}_{i,n}(k) - \zeta_{i,n}\sum_{l=1}^{\rho}\bar{\phi}_{i,n}(\omega_{i,n}(k-l+1)) \\ &\quad \times [\chi_{i,n}(k-l+1) + \chi_{i,o}(k-l+1)]\end{aligned} \tag{28}$$

where $\zeta_{i,n} > 0$ is the convergence factor.

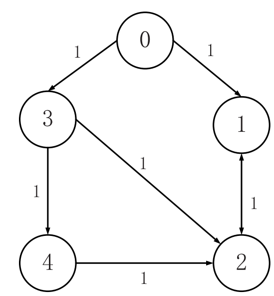


Fig. 2. Communication topology.

C. Stability Analysis

The main theorem of this article is as follows.

Theorem 1: Considering the feasible virtual control signals $\breve{\varrho}_{i,1}(k)$ in (12), $\breve{\varrho}_{i,m}(k)$ in (19), the multigradient recursive updating laws $\hat{\bar{\varpi}}_{i,o}(k+1)$ in (9), $\hat{\bar{\varpi}}_{i,1}(k+1)$ in (16), $\hat{\bar{\varpi}}_{i,m}(k+1)$ in (21), $\hat{\bar{\varpi}}_{i,n}(k+1)$ in (28), and controller $u_i(k)$ in (26), all signals in the closed-loop MASs (1) are SGUUB.

The proof of Theorem 1 can be seen in the Appendix.

IV. Simulation Results

In this section, two simulation examples are addressed to test the effectiveness of the designed algorithm.

We suppose that the nonlinear MASs consist of a leader and four followers. Fig. 2 shows the communication topology with **0** being the leader and 1–4 being four followers.

Example 1: Choose the numerical example as follows:

$$\begin{aligned}s_{i,1}(k+1) &= \frac{0.02s_{i,1}^2(k)}{0.8 + s_{i,1}^2(k)} + \tau_{i,1}s_{i,2}(k) \\ s_{i,2}(k+1) &= \frac{s_{i,2}^2(k)}{0.5 + s_{i,1}^2(k) + s_{i,2}^2(k)} + \tau_{i,2}\nu_i(k) + d_i(k) \\ y_i &= \breve{\zeta}_i(k)s_{i,1}(k) + \iota_i(k)\end{aligned}$$

where

$$d_i(k) = 1.09\cos(s_{i,1}(k))\cos(0.064k)$$

with $i = 1, 2, 3$, and 4. In addition, the desired signal is selected as $y_d(k) = (3/20) + (3/10)\cos(-25\pi/14 + 50kT\pi/299)$.

We define the parameters $s_1(0) = [0.4, 0.2]^T$, $s_2(0) = [0.4, 0.2]^T$, $s_3(0) = [0.4, 0.2]^T$, $s_4(0) = [0.4, 0.2]^T$, $\zeta_{i,o} = 0.0006$, $\zeta_{i,1} = 0.026$, $\zeta_{i,2} = 0.0344$, $\tau_{i,j} = 0.245$, $\breve{\sigma}_i = 0.00035$, $\breve{\zeta}_1(k) = 1$, $\iota_1(k) = 0.3$, $\breve{\zeta}_2(k) = 0.9$, $\iota_2(k) = 0$, $\breve{\zeta}_3(k) = 1$, $\iota_3(k) = -e^{-s_{3,1}(k)}$, $\breve{\zeta}_4(k) = 1$, and $\iota_4(k) = 1$ with $i = 1, 2, 3$, and 4 and $j = 1$ and 2.

The numerical simulation results are given in Figs. 3–6. Fig. 3 shows that the outputs of followers can converge to a neighborhood of the leader's output. The responses of controllers and input signals of MASs are shown in Fig. 4. Fig. 5 shows the event triggering time of u_1–u_4, and it can be seen that the event-triggered controllers can save communication resources. In addition, Fig. 6 shows the trajectories of $\hat{\bar{\varpi}}_{1,o}$, $\hat{\bar{\varpi}}_{1,1}$, and $\hat{\bar{\varpi}}_{1,2}$.

Example 2: To further verify the effectiveness of the proposed control method, a stirred tank reactor practical system

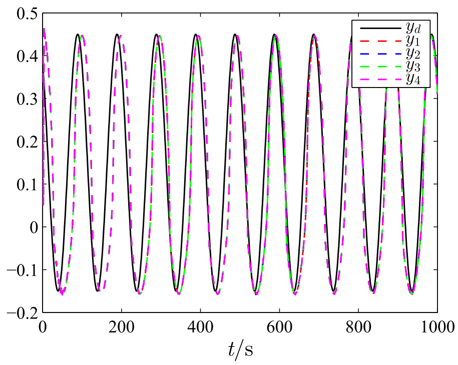


Fig. 3. Trajectories of followers y_i and leader y_d with $i = 1, 2, 3$, and 4 in Example 1.

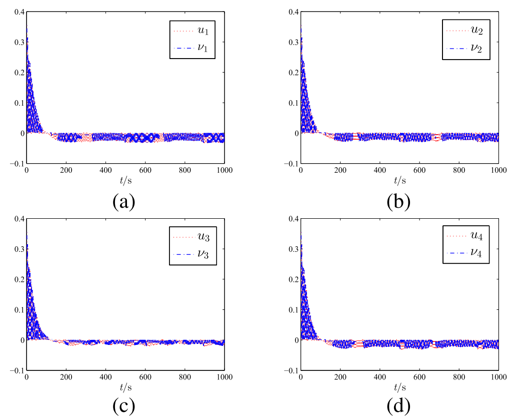


Fig. 4. Controllers and input signals of MASs in Example 1. (a) Trajectories of controller u_1 and input signal ν_1. (b) Trajectories of controller u_2 and input signal ν_2. (c) Trajectories of controller u_3 and input signal ν_3. (d) Trajectories of controller u_4 and input signal ν_4.

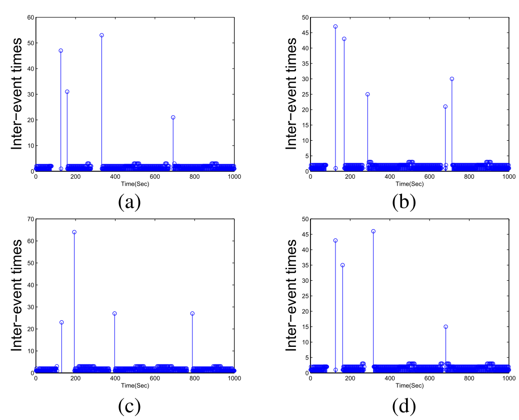


Fig. 5. Triggering time of u_i with $i = 1, 2, 3$, and 4 in Example 1. (a) Triggering time of u_1. (b) Triggering time of u_2. (c) Triggering time of u_3. (d) Triggering time of u_4.

is considered. The same topology is selected as Example 1. In addition, each parameter's physical meaning is shown in Table I, and the system [60] is described as follows:

$$s_{i,1}(k+1) = s_{i,1}(k) + T(-s_{i,1}(k) + \Gamma_{i,1}(k))$$
$$s_{i,2}(k+1) = s_{i,2}(k) + T(-s_{i,2} + \check{\mu}_{i,1}\Gamma_{i,1}(k) - \Gamma_{i,2}(k)) + d_i(k), \quad i = 1, 2, 3, 4$$

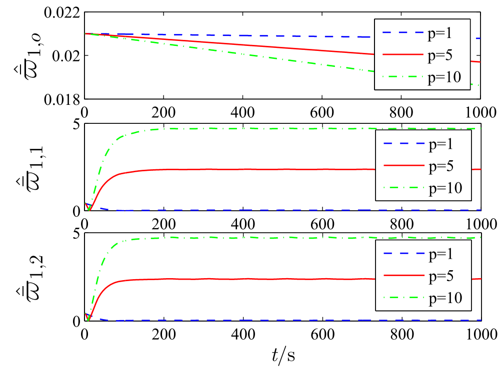


Fig. 6. Trajectories of $\hat{\breve{\varpi}}_{1,o}$, $\hat{\breve{\varpi}}_{1,1}$ and $\hat{\breve{\varpi}}_{1,2}$ in Example 1.

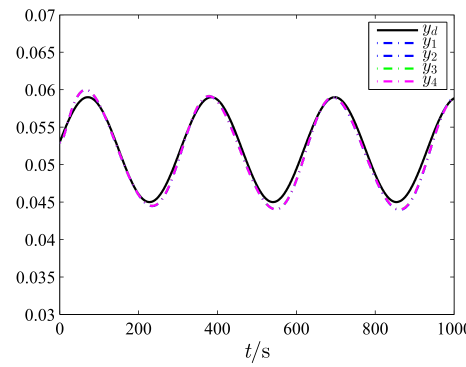


Fig. 7. Trajectories of followers y_i and leader y_d with $i = 1, 2, 3$, and 4 in Example 2.

TABLE I
PARAMETERS OF SIMULATION

$\acute{\sigma}_{i,1} = 0.04$	the Damkohler number
$\acute{\sigma}_{i,2} = 25.2$	the dimensionless cooling rate
$\check{\mu}_{i,1} = 21$	the heat of reaction for dimensionless
$s_{i,1}(k)$	the dimensionless concentration at time k
$s_{i,2}(k)$	the dimensionless temperature
$d_i(k)$	the disturbance
T	the sampling period

with

$$\Gamma_{i,1}(k) = \acute{\sigma}_{i,1}(1.23 - 2.2s_{i,1})e^{\frac{\breve{\varpi}_i s_{i,2}(k)}{\breve{\varpi}_i + s_{i,2}(k)}}$$
$$\Gamma_{i,2}(k) = \acute{\sigma}_{i,2}(s_{i,2}(k) - \nu_i(k))$$

where $\breve{\varpi}_i = 28.5$ and $d_i(k) = (2/25)\cos((1/50)k) \cos(s_{i,1}(k))$. Moreover, the signal of the leader is $y_d(k) = (7/1000)\cos((97/20) + (17/2660)k\pi) + (9/200)$.

We define the parameters $s_1(0) = [0.053, 0.053]^T$, $s_2(0) = [0.053, 0.053]^T$ and $s_3(0) = [0.053, 0.053]^T$, $s_4(0) = [0.053, 0.053]^T$, and the remaining parameters are the same as Example 1.

The practical simulation results are shown in Figs. 7–10. Fig. 7 verifies the tracking control performance of the MASs. Fig. 8 shows the responses of controllers and input signals of MASs. Fig. 9 describes the event triggering time of u_1–u_4. Fig. 10 shows the trajectories of multigradient recursive updating laws. From Figs. 3 to 10, it can be seen that all signals of the discrete-time closed-loop MASs are SGUUB.

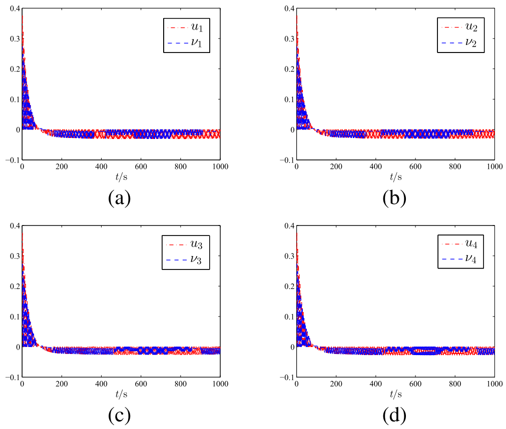


Fig. 8. Controllers and input signals of MASs in Example 2. (a) Trajectories of controller u_1 and input signal ν_1. (b) Trajectories of controller u_2 and input signal ν_2. (c) Trajectories of controller u_3 and input signal ν_3. (d) Trajectories of controller u_4 and input signal ν_4.

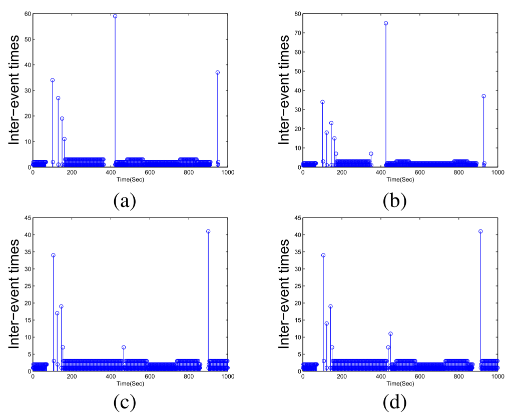


Fig. 9. Triggering time of u_i with $i = 1, 2, 3$, and 4 in Example 2. (a) Triggering time of u_1. (b) Triggering time of u_2. (c) Triggering time of u_3. (d) Triggering time of u_4.

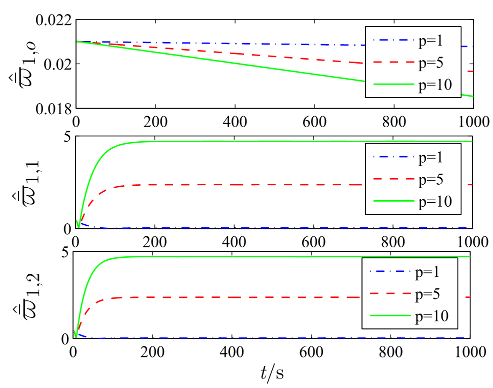


Fig. 10. Trajectories of $\hat{\bar{\varpi}}_{1,o}$, $\hat{\bar{\varpi}}_{1,1}$ and $\hat{\bar{\varpi}}_{1,2}$ in Example 2.

Remark 3: To obtain the desired control performance, select and adjust these parameters under the premise of meeting the design conditions by our experience. In addition, these parameters should be selected according to different system models.

V. Conclusion

In this article, a fault-tolerant adaptive multigradient recursive RL event-triggered consensus control strategy has been presented for strict-feedback discrete-time MASs. The multigradient recursive RL algorithm has been used to avoid the local optimal problem existed in the gradient descent scheme. Then, a new event-triggered relative threshold control strategy with compensation operator has been addressed to improve the utilization of communication resources and weaken the negative impact of the event-triggered threshold on tracking accuracy and stability of closed-loop MASs. A distributed fault-tolerant control method has been used, which can effectively reduce the number of online estimation parameters. Moreover, by combining the direct adaptive control method and the multigradient recursive RL algorithm with LLPs, the computational burden has been effectively reduced. It has been proved that all signals of the closed-loop MASs are SGUUB. At last, the effectiveness of the proposed control scheme has been testified by two simulation examples. Switching threshold event-triggered multigradient recursive RL control problem of discrete-time MASs will be an interesting topic for our further work.

Appendix

This section will give the proof of Theorem 1.

Step 1: The following Lyapunov function candidate is considered:

$$V_{i,1}(k) = V_{i,a1}(k) + V_{i,b1}(k) + V_{i,c1}(k) + V_{i,d1}(k)$$

with

$$\begin{aligned}
V_{i,a1}(k) &= \frac{\alpha_{i,a1}}{3(\theta_i + \varsigma_{i,\mathbf{0}})^2}\xi_{i,1}^2(k),\\
V_{i,b1}(k) &= \frac{\alpha_{i,b1}}{\zeta_{i,1}}\sum_{\check{j}=0}^{n-1}\tilde{\bar{\varpi}}_{i,1}^2(k_1+\check{j}),\\
V_{i,c1}(k) &= \frac{\alpha_{i,o}}{\zeta_{i,o}}\tilde{\bar{\varpi}}_{i,o}^2(k),\\
V_{i,d1}(k) &= 2\alpha_{i,o}\sum_{l=1}^{\rho}e_{i,o}^2(k-l)
\end{aligned}$$

where $\alpha_{i,a1} > 0$, $\alpha_{i,b1} > 0$ and $\alpha_{i,o} > 0$ are constants, and $e_{i,o}(k-l) = \tilde{\bar{\varpi}}_{i,o}(k)\bar{\phi}_{i,o}(k-l)$ with $\tilde{\bar{\varpi}}_{i,o}(k) = \hat{\bar{\varpi}}_{i,o}(k) - \bar{\varpi}_{i,o}$ being the estimation error.

In the light of Cauchy–Schwarz inequality, it yields

$$\begin{aligned}
\triangle V_{i,a1}(k) \le\ & \alpha_{i,a1}\tau_{i,1}^2\xi_{i,2}^2(k) + \alpha_{i,a1}\tau_{i,1}^2 e_{i,1}^2(k_1)\\
&+ \alpha_{i,a1}\tau_{i,1}^2\epsilon_{i,1}^T(k_1)\epsilon_{i,1}(k_1) - \frac{\alpha_{i,a1}}{3(\theta_i+\varsigma_{i,\mathbf{0}})^2}\xi_{i,1}^2(k)\\
\le\ & \alpha_{i,a1}\tau_{i,1}^2\xi_{i,2}^2(k) + \alpha_{i,a1}\tau_{i,1}^2 e_{i,1}^2(k_1)\\
&+ \alpha_{i,a1}\tau_{i,1}^2\bar{\epsilon}_{i,1} - \frac{\alpha_{i,a1}}{3(\theta_i+\varsigma_{i,\mathbf{0}})^2}\xi_{i,1}^2(k),\\
\triangle V_{i,b1}(k) \le\ & -\alpha_{i,b1}(1-\rho\beta\zeta_{i,1})\\
&\times\sum_{l=1}^{\rho}(\chi_{i,1}(k_1-l+1) + \chi_{i,o}(k-l+1))^2\\
&- \alpha_{i,b1}\sum_{l=1}^{\rho}e_{i,1}^2(k_1-l+1)
\end{aligned}$$

$$+ 2\alpha_{i,b1}\sum_{l=1}^{\rho} e_{i,o}^2(k-l+1) + 2\rho\beta\alpha_{i,b1}[\bar{\varpi}_{i,o} + \bar{\varpi}_{i,1}]^2,$$

$$\triangle V_{i,c1}(k) \le -\alpha_{i,o}(1-\rho\beta\zeta_{i,o}\check{o}_i^2) \times \sum_{l=1}^{\rho}(\check{o}_i\chi_{i,o}(k-l+1) + \Upsilon_i(k-l+1) - \chi_{i,o}(k-l))^2 - \alpha_{i,o}\check{o}_i^2\sum_{l=1}^{\rho} e_{i,o}^2(k-l+1) + 2\alpha_{i,o}\sum_{l=1}^{\rho} e_{i,o}^2(k-l) + 2\rho\alpha_{i,o} \times [\bar{\varpi}_{i,o}\beta(1+\check{o}_i)+1]^2$$

$$\triangle V_{i,d1}(k) = 2\alpha_{i,o}\sum_{l=1}^{\rho} e_{i,o}^2(k-l+1) - 2\alpha_{i,o} \times \sum_{l=1}^{\rho} e_{i,o}^2(k-l)$$

where $e_{i,1}(k_1) = \tilde{\bar{\varpi}}_{i,1}(k_1)\bar{\phi}_{i,1}(\omega_{i,1}(k_1))$ and $\epsilon_{i,1}^T(k_1)\epsilon_{i,1}(k_1) \le (\bar{\varepsilon}_{i,1} + 2\sqrt{\beta}\bar{\varpi}_{i,1})^2 = \bar{\epsilon}_{i,1}$ with $\bar{\epsilon}_{i,1} > 0$ being a constant and $\varepsilon_{i,1}(k) \le \bar{\varepsilon}_{i,1}$. Then, one has

$$\begin{aligned}\triangle V_{i,1}(k) &\le -\alpha_{i,b1}(1-\rho\beta\zeta_{i,1}) \times \sum_{l=1}^{\rho}(\chi_{i,1}(k_1-l+1) + \chi_{i,o}(k-l+1))^2 \\ &\quad - \alpha_{i,o}\big(1-\rho\beta\zeta_{i,o}\check{o}_i^2\big) \times \sum_{l=1}^{\rho}(\check{o}_i\chi_{i,o}(k-l+1) + \Upsilon_i(k-l+1) - \chi_{i,o}(k-l))^2 \\ &\quad - \frac{\alpha_{i,a1}}{3(\theta_i+\varsigma_{i,\mathbf{0}})^2}\xi_{i,1}^2(k) - \big(\alpha_{i,b1} - \alpha_{i,a1}\tau_{i,1}^2\big)e_{i,1}^2(k_1) \\ &\quad - \alpha_{i,b1}\sum_{l=2}^{\rho} e_{i,1}^2(k_1-l+1) - \big(\alpha_{i,o}\check{o}_i^2 - 2\alpha_{i,b1} - 2\alpha_{i,o}\big) \\ &\quad \times \sum_{l=1}^{\rho} e_{i,o}^2(k-l+1) + \alpha_{i,a1}\tau_{i,1}^2\xi_{i,2}^2(k) + r_{i,1},\end{aligned} \tag{29}$$

where $r_{i,1} = \alpha_{i,a1}\tau_{i,1}^2\bar{\epsilon}_{i,1} + 2\rho\beta\alpha_{i,b1}[\bar{\varpi}_{i,o} + \bar{\varpi}_{i,1}]^2 + 2\rho\alpha_{i,o}[\bar{\varpi}_{i,o}\beta(1+\check{o}_i)+1]^2$.

Step m ($2 \le m \le n-1$): The following Lyapunov function candidate is considered:

$$V_{i,m}(k) = V_{i,am}(k) + V_{i,bm}(k)$$

with

$$V_{i,am}(k) = \frac{\alpha_{i,am}}{3}\xi_{i,m}^2(k)$$
$$V_{i,bm}(k) = \frac{\alpha_{i,bm}}{\zeta_{i,m}}\sum_{\check{j}=0}^{n-m}\tilde{\bar{\varpi}}_{i,m}^2(k_m+\check{j})$$

where $\alpha_{i,am}$ and $\alpha_{i,bm}$ are positive constants.

In the light of the forward difference, it yields

$$\begin{aligned}\triangle V_{i,am}(k) &\le \alpha_{i,am}\tau_{i,m}^2\xi_{i,m+1}^2(k) + \alpha_{i,am}\tau_{i,m}^2 e_{i,m}^2(k_m) \\ &\quad + \alpha_{i,am}\tau_{i,m}^2\epsilon_{i,m}^T(k_m)\epsilon_{i,m}(k_m) - \frac{\alpha_{i,am}}{3}\xi_{i,m}^2(k) \\ &\le \alpha_{i,am}\tau_{i,m}^2\xi_{i,m+1}^2(k) + \alpha_{i,am}\tau_{i,m}^2 \\ &\quad \times e_{i,m}^2(k_m) + \alpha_{i,am}\tau_{i,m}^2\bar{\epsilon}_{i,m} - \frac{\alpha_{i,am}}{3}\xi_{i,m}^2(k)\end{aligned}$$

$$\begin{aligned}\triangle V_{i,bm}(k) &\le -\alpha_{i,bm}(1-\rho\beta\zeta_{i,m}) \times \sum_{l=1}^{\rho}(\chi_{i,m}(k_m-l+1) + \chi_{i,o}(k-l+1))^2 \\ &\quad - \alpha_{i,bm}\sum_{l=1}^{\rho} e_{i,m}^2(k_m-l+1) + 2\alpha_{i,bm}\sum_{l=1}^{\rho} e_{i,o}^2(k-l+1) \\ &\quad + 2\rho\beta\alpha_{i,bm}[\bar{\varpi}_{i,o} + \bar{\varpi}_{i,m}]^2,\end{aligned}$$

where $e_{i,m}(k_m) = \tilde{\bar{\varpi}}_{i,m}(k_m)\bar{\phi}_{i,m}(\omega_{i,m}(k_m))$ and $\epsilon_{i,m}^T(k_m)\epsilon_{i,m}(k_m) \le (\bar{\varepsilon}_{i,m} + 2\sqrt{\beta}\bar{\varpi}_{i,m})^2 = \bar{\epsilon}_{i,m}$ with $\bar{\epsilon}_{i,m} > 0$ being a constant and $\varepsilon_{i,m} \le \bar{\varepsilon}_{i,m}$. Then, it yields

$$\begin{aligned}\triangle V_{i,m}(k) &\le -\alpha_{i,bm}(1-\rho\beta\zeta_{i,m}) \times \sum_{l=1}^{\rho}(\chi_{i,m}(k_m-l+1) + \chi_{i,o}(k-l+1))^2 \\ &\quad - \frac{\alpha_{i,am}}{3}\xi_{i,m}^2(k) + 2\alpha_{i,bm}\sum_{l=1}^{\rho} e_{i,o}^2(k-l+1) \\ &\quad - \big(\alpha_{i,bm} - \alpha_{i,am}\tau_{i,m}^2\big)e_{i,m}^2(k_m) \\ &\quad - \alpha_{i,bm}\sum_{l=2}^{\rho} e_{i,m}^2(k_m-l+1) \\ &\quad + r_{i,m} + \alpha_{i,am}\tau_{i,m}^2\xi_{i,m+1}^2(k),\end{aligned} \tag{30}$$

where $r_{i,m} = \alpha_{i,am}\tau_{i,m}^2\bar{\epsilon}_{i,m} + 2\rho\beta\alpha_{i,bm}[\bar{\varpi}_{i,o} + \bar{\varpi}_{i,m}]^2$.

Step n: The following Lyapunov function candidate is considered:

$$V_{i,n}(k) = V_{i,an}(k) + V_{i,bn}(k)$$

with

$$V_{i,an}(k) = \frac{\alpha_{i,an}}{4}\xi_{i,n}^2(k)$$
$$V_{i,bn}(k) = \frac{\alpha_{i,bn}}{\zeta_{i,n}}\tilde{\bar{\varpi}}_{i,n}^2(k)$$

where $\alpha_{i,an}$ and $\alpha_{i,bn}$ are positive constants.

On the basis of the forward difference, it yields

$$\begin{aligned}\triangle V_{i,an}(k) &\le \alpha_{i,an}\tau_{i,n}^2(1+\varphi_i(k)\check{\eta}_i)^2 e_{i,n}^2(k) \\ &\quad + \alpha_{i,an}\tau_{i,n}^2\epsilon_{i,n}^T(k)\epsilon_{i,n}(k) + \alpha_{i,an}d_i^2(k) \\ &\quad + \alpha_{i,an}\tau_{i,n}^2\pi_i^2(k) - \frac{\alpha_{i,an}}{4}\xi_{i,n}^2(k) \\ &\le \alpha_{i,an}\tau_{i,n}^2(1+\varphi_i(k)\check{\eta}_i)^2 e_{i,n}^2(k) \\ &\quad + \alpha_{i,an}\tau_{i,n}^2\bar{\epsilon}_{i,n} + \alpha_{i,an}\bar{d}^2 \\ &\quad + \alpha_{i,an}\tau_{i,n}^2\pi_m^2 - \frac{\alpha_{i,an}}{4}\xi_{i,n}^2(k),\end{aligned}$$

$$\triangle V_{i,bn}(k) \le -\alpha_{i,bn}(1-\rho\beta\zeta_{i,n})$$

$$\times\sum_{l=1}^{\rho}(\chi_{i,n}(k-l+1)+\chi_{i,o}(k-l+1))^2$$
$$-\alpha_{i,bn}\sum_{l=1}^{\rho}e_{i,n}^2(k-l+1)$$
$$+2\alpha_{i,bn}\sum_{l=1}^{\rho}e_{i,o}^2(k-l+1)$$
$$+2\rho\beta\alpha_{i,bn}[\bar{\varpi}_{i,o}+\bar{\varpi}_{i,n}]^2$$

where $e_{i,n}(k)=\tilde{\bar{\varpi}}_{i,n}(k)\bar{\phi}_{i,n}(\omega_{i,n}(k))$ and $\epsilon_{i,n}^T(k)\epsilon_{i,n}(k)\le(1+\varphi_i(k)\check{\eta}_i)^2(\bar{\varepsilon}_{i,n}+2\sqrt{\beta}\bar{\varpi}_{i,n})^2=\bar{\epsilon}_{i,n}$ with $\bar{\epsilon}_{i,n}>0$ being a constant and $\varepsilon_{i,n}\le\bar{\varepsilon}_{i,n}$. In addition, by using Lemma 1, it yields that $\pi_i(k)\le\pi_m$ with $\pi_m>0$ being a constant. Then, it gives

$$\begin{aligned}\triangle V_{i,n}(k)\le&-\alpha_{i,bn}(1-\rho\beta\zeta_{i,n})\\&\times\sum_{l=1}^{\rho}(\chi_{i,n}(k-l+1)+\chi_{i,o}(k-l+1))^2\\&-\frac{\alpha_{i,an}}{4}\xi_{i,n}^2(k)+2\alpha_{i,bn}\sum_{l=1}^{\rho}e_{i,o}^2(k-l+1)\\&-\left(\alpha_{i,bn}-\alpha_{i,an}\tau_{i,n}^2(1+\varphi_i(k)\check{\eta}_i)^2\right)e_{i,n}^2(k)\\&-\alpha_{i,bn}\sum_{l=2}^{\rho}e_{i,n}^2(k-l+1)+r_{i,n}\end{aligned}\tag{31}$$

where $r_{i,n}=\alpha_{i,an}\tau_{i,n}^2\bar{\epsilon}_{i,n}+\alpha_{i,an}\bar{d}^2+\alpha_{i,an}\tau_{i,n}^2\pi_m^2+2\rho\beta\alpha_{i,bn}[\bar{\varpi}_{i,o}+\bar{\varpi}_{i,n}]^2$. Then, we let

$$\begin{aligned}V_i(k)=&\sum_{\psi=1}^{n}V_{i,\psi}\\=&\sum_{\psi=1}^{n}\frac{\alpha_{i,b\psi}}{\zeta_{i,\psi}}\sum_{\check{j}=0}^{n-\psi}\tilde{\bar{\varpi}}_{i,\psi}^2(k_\psi+\check{j})+\frac{\alpha_{i,a1}}{3(\theta_i+\varsigma_{i,\mathbf{0}})^2}\\&\times\xi_{i,1}^2(k)+\sum_{\psi=2}^{n-1}\frac{\alpha_{i,a\psi}}{3}\xi_{i,\psi}^2(k)+\frac{\alpha_{i,an}}{4}\xi_{i,n}^2(k)\\&+\frac{\alpha_{i,o}}{\zeta_{i,o}}\tilde{\bar{\varpi}}_{i,o}^2(k)+2\alpha_{i,o}\sum_{l=1}^{\rho}e_{i,o}^2(k-l)\end{aligned}$$

where $k_n=k$.

Consider the Lyapunov function as

$$V(k)=\sum_{i=1}^{N}V_i(k).$$

By combining (29)–(31), we can obtain

$$\begin{aligned}&\triangle V(k)\\&\le-\sum_{i=1}^{N}\sum_{\psi=1}^{n}\alpha_{i,b\psi}(1-\rho\beta\zeta_{i,\psi})\\&\quad\times\sum_{l=1}^{\rho}(\chi_{i,\psi}(k_\psi-l+1)+\chi_{i,o}(k-l+1))^2\\&\quad-\sum_{i=1}^{N}\alpha_{i,o}\left(1-\rho\beta\zeta_{i,o}\check{o}_i^2\right)\\&\quad\times\sum_{l=1}^{\rho}(\check{o}_i\chi_{i,o}(k-l+1)+\Upsilon_i(k-l+1)-\chi_{i,o}(k-l))^2\\&\quad-\sum_{i=1}^{N}\frac{\alpha_{i,a1}}{3(\theta_i+\varsigma_{i,\mathbf{0}})^2}\times\xi_{i,1}^2(k)-\sum_{i=1}^{N}\sum_{\psi=2}^{n-1}\frac{\alpha_{i,a\psi}}{3}\xi_{i,\psi}^2(k)\\&\quad-\sum_{i=1}^{N}\frac{\alpha_{i,an}}{4}\xi_{i,n}^2(k)-\sum_{i=1}^{N}\sum_{\psi=1}^{n-1}\left(\alpha_{i,b\psi}-\alpha_{i,a\psi}\tau_{i,\psi}^2\right)e_{i,\psi}^2(k_\psi)\\&\quad-\sum_{i=1}^{N}\left(\alpha_{i,bn}-\alpha_{i,an}\tau_{i,n}^2(1+\varphi_i(k)\check{\eta}_i)^2\right)e_{i,n}^2(k)\\&\quad-\sum_{i=1}^{N}\sum_{\psi=1}^{n}\alpha_{i,b\psi}\sum_{l=2}^{\rho}e_{i,\psi}^2(k_\psi-l+1)\\&\quad-\sum_{i=1}^{N}\left(\alpha_{i,o}\check{o}_i^2-2\alpha_{i,b1}-2\alpha_{i,o}\right)\sum_{l=1}^{\rho}e_{i,o}^2(k-l+1)\\&\quad+\sum_{i=1}^{N}\sum_{\psi=2}^{n}2\alpha_{i,b\psi}\times\sum_{l=1}^{\rho}e_{i,o}^2(k-l+1)\\&\quad+\sum_{i=1}^{N}\sum_{\psi=1}^{n-1}\alpha_{i,a\psi}\tau_{i,\psi}^2\xi_{i,\psi+1}^2(k)+r\end{aligned}$$

where $r=\sum_{i=1}^{N}r_i$ with $r_i=\sum_{\psi=1}^{n}r_{i,\psi}$.

Choose the proper parameters such that $0<\zeta_{i,o}<(1/(\rho\beta\check{o}_i^2))$, $0<\zeta_{i,\psi}<(1/\rho\beta)$, where $\psi=1,\ldots,n$. Hence, it results in

$$\begin{aligned}\triangle V(k)\le&-\sum_{i=1}^{N}\frac{\alpha_{i,a1}}{3(\theta_i+\varsigma_{i,\mathbf{0}})^2}\xi_{i,1}^2(k)\\&-\sum_{i=1}^{N}\sum_{\psi=2}^{n-1}\left(\frac{\alpha_{i,a\psi}}{3}-\alpha_{i,a(\psi-1)}\tau_{i,(\psi-1)}^2\right)\xi_{i,\psi}^2(k)\\&-\sum_{i=1}^{N}\left(\frac{\alpha_{i,an}}{4}-\alpha_{i,a(n-1)}\tau_{i,(n-1)}^2\right)\xi_{i,n}^2(k)\\&-\sum_{i=1}^{N}\sum_{\psi=1}^{n-1}\left(\alpha_{i,b\psi}-\alpha_{i,a\psi}\tau_{i,\psi}^2\right)e_{i,\psi}^2(k_\psi)\\&-\sum_{i=1}^{N}\left(\alpha_{i,bn}-\alpha_{i,an}\tau_{i,n}^2(1+\varphi_i(k)\check{\eta}_i)^2\right)e_{i,n}^2(k)\\&-\sum_{i=1}^{N}\sum_{\psi=1}^{n}\alpha_{i,b\psi}\sum_{l=2}^{\rho}e_{i,\psi}^2(k_\psi-l+1)\\&-\sum_{i=1}^{N}\left(\alpha_{i,o}\check{o}_i^2-2\alpha_{i,b1}-2\alpha_{i,o}-\sum_{\psi=2}^{n}2\alpha_{i,b\psi}\right)\\&\times\sum_{l=1}^{\rho}e_{i,o}^2(k-l+1)+r.\end{aligned}$$

Clearly, the parameters satisfy

$$\begin{aligned}\alpha_{i,a\psi}&>3\alpha_{i,a(\psi-1)}\tau_{i,(\psi-1)}^2,\quad\psi=2,\ldots,n-1\\\alpha_{i,an}&>4\alpha_{i,a(n-1)}\tau_{i,(n-1)}^2\\\alpha_{i,b\psi}&>\alpha_{i,a\psi}\tau_{i,\psi}^2,\quad\psi=1,\ldots,n-1\end{aligned}$$

$$\alpha_{i,bn} > \alpha_{i,an}\tau_{i,n}^2(1+\varphi_i(k)\check{\eta}_i)^2$$
$$\alpha_{i,o}\check{\delta}_i^2 > 2\alpha_{i,b1} + 2\alpha_{i,o} + \sum_{\psi=2}^{n} 2\alpha_{i,b\psi}.$$

Based on [32], we can get $\triangle V(k) < 0$ from the following conditions:

$$|\xi_{i,1}(k)| > \frac{\sqrt{r}}{\sqrt{\alpha_{i,a1}/(3(\theta_i + \varsigma_{i,\mathbf{0}})^2)}}$$

or

$$|\xi_{i,\psi}(k)| > \frac{\sqrt{r}}{\sqrt{\Phi_1}}, \quad \psi = 2, \ldots, n-1$$

or

$$|\xi_{i,n}(k)| > \frac{\sqrt{r}}{\sqrt{\alpha_{i,an}/4 - \alpha_{i,a(n-1)}\tau_{i,(n-1)}^2}}$$

or

$$|e_{i,\psi}(k)| > \frac{\sqrt{r}}{\sqrt{\alpha_{i,b\psi} - \alpha_{i,a\psi}\tau_{i,\psi}^2}}, \quad \psi = 1, \ldots, n-1$$

or

$$|e_{i,n}(k)| > \frac{\sqrt{r}}{\sqrt{\alpha_{i,bn} - \alpha_{i,an}\tau_{i,n}^2(1+\varphi_i(k)\check{\eta}_i)^2}}$$

or

$$\left|\sum_{l=2}^{\rho} e_{i,\psi}^2(k_\psi - l + 1)\right| > \frac{\sqrt{r}}{\sqrt{\alpha_{i,b\psi}}}, \quad \psi = 1, \ldots, n$$

or

$$\left|\sum_{l=1}^{\rho} e_{i,o}^2(k - l + 1)\right| > \frac{\sqrt{r}}{\sqrt{\Phi_2}}$$

where $\Phi_1 = \alpha_{i,a\psi}/3 - \alpha_{i,a(\psi-1)}\tau_{i,(\psi-1)}^2$, $\Phi_2 = \alpha_{i,o}\check{\delta}_i^2 - 2\alpha_{i,b1} - 2\alpha_{i,o} - \sum_{\psi=2}^{n} 2\alpha_{i,b\psi}$. Then, it means that all signals of closed-loop MASs are SGUUB.

Remark 4: Compared with the results in [30]–[32], the multigradient recursive RL algorithm with LLPs is used in this article, which can avoid the local optimal issue and decrease online estimation time. In addition, different from the traditional relative threshold event-triggered strategy [35]–[41], a compensation operator is introduced in this article, which can reduce the communication burden and weaken the negative impact of the event-triggered threshold on tracking accuracy and stability of closed-loop MASs.

REFERENCES

[1] V. Vahidpour, A. Rastegarnia, A. Khalili, and S. Sanei, "Partial diffusion Kalman filtering for distributed state estimation in multiagent networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 12, pp. 3839–3846, Dec. 2019.

[2] Q. Shen, P. Shi, R. K. Agarwal, and Y. Shi, "Adaptive neural network-based filter design for nonlinear systems with multiple constraints," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jul. 28, 2020, doi: 10.1109/TNNLS.2020.3009391.

[3] Y. Yang, Y. Li, and D. Yue, "Event-trigger-based consensus secure control of linear multi-agent systems under DoS attacks over multiple transmission channels," *Sci. China Inf. Sci.*, vol. 63, no. 5, May 2020, Art. no. 150208.

[4] Z. Li, L. Gao, W. Chen, and Y. Xu, "Distributed adaptive cooperative tracking of uncertain nonlinear fractional-order multi-agent systems," *IEEE/CAA J. Automatica Sinica*, vol. 7, no. 1, pp. 292–300, Jan. 2020.

[5] H. Ren, H. R. Karimi, R. Lu, and Y. Wu, "Synchronization of network systems via aperiodic sampled-data control with constant delay and application to unmanned ground vehicles," *IEEE Trans. Ind. Electron.*, vol. 67, no. 6, pp. 4980–4990, Jun. 2020.

[6] Y. Su, Q. Wang, and C. Sun, "Self-triggered consensus control for linear multi-agent systems with input saturation," *IEEE/CAA J. Automatica Sinica*, vol. 7, no. 1, pp. 150–157, Jan. 2020.

[7] B. Fan, S. Guo, J. Peng, Q. Yang, W. Liu, and L. Liu, "A consensus-based algorithm for power sharing and voltage regulation in DC microgrids," *IEEE Trans. Ind. Informat.*, vol. 16, no. 6, pp. 3987–3996, Jun. 2020.

[8] C. Deng and G.-H. Yang, "Distributed adaptive fault-tolerant control approach to cooperative output regulation for linear multi-agent systems," *Automatica*, vol. 103, pp. 62–68, May 2019.

[9] Y. Liu, D. Yao, H. Li, and R. Lu, "Distributed cooperative compound tracking control for a platoon of vehicles with adaptive NN," *IEEE Trans. Cybern.*, early access, Jan. 11, 2021, doi: 10.1109/TCYB.2020.3044880.

[10] G. Dong, H. Li, H. Ma, and R. Lu, "Finite-time consensus tracking neural network FTC of multi-agent systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 2, pp. 653–662, Feb. 2021.

[11] Y. Wu, H. Liang, Y. Zhang, and C. K. Ahn, "Cooperative adaptive dynamic surface control for a class of high-order stochastic nonlinear multiagent systems," *IEEE Trans. Cybern.*, early access, May 12, 2020, doi: 10.1109/TCYB.2020.2986332.

[12] H. He, S. Chen, K. Li, and X. Xu, "Incremental learning from stream data," *IEEE Trans. Neural Netw.*, vol. 22, no. 12, pp. 1901–1914, Dec. 2011.

[13] Z. Ni, H. He, and J. Wen, "Adaptive learning in tracking control based on the dual critic network design," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, no. 6, pp. 913–928, Jun. 2013.

[14] N. Wang, Y. Gao, H. Zhao, and C. K. Ahn, "Reinforcement learning-based optimal tracking control of an unknown unmanned surface vehicle," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Aug. 3, 2020, doi: 10.1109/TNNLS.2020.3009214.

[15] X.-J. Li and X.-Y. Shen, "A data-driven attack detection approach for DC servo motor systems based on mixed optimization strategy," *IEEE Trans. Ind. Informat.*, vol. 16, no. 9, pp. 5806–5813, Sep. 2020.

[16] M. Liu, Y. Wan, F. L. Lewis, and V. G. Lopez, "Adaptive optimal control for stochastic multiplayer differential games using on-policy and off-policy reinforcement learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 12, pp. 5522–5533, Dec. 2020.

[17] Y.-J. Liu, S. Li, S. Tong, and C. L. P. Chen, "Adaptive reinforcement learning control based on neural approximation for nonlinear discrete-time systems with unknown nonaffine dead-zone input," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 1, pp. 295–305, Jan. 2019.

[18] X. Guo, W. Yan, and R. Cui, "Event-triggered reinforcement learning-based adaptive tracking control for completely unknown continuous-time nonlinear systems," *IEEE Trans. Cybern.*, vol. 50, no. 7, pp. 3231–3242, Jul. 2020.

[19] Y. Yang, Z. Guo, H. Xiong, D.-W. Ding, Y. Yin, and D. C. Wunsch, "Data-driven robust control of discrete-time uncertain linear systems via off-policy reinforcement learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 12, pp. 3735–3747, Dec. 2019.

[20] M. V. Butz, D. E. Goldberg, and P. L. Lanzi, "Gradient descent methods in learning classifier systems: Improving XCS performance in multistep problems," *IEEE Trans. Evol. Comput.*, vol. 9, no. 5, pp. 452–473, Oct. 2005.

[21] X. Dai, C. K. Li, and A. B. Rad, "An approach to tune fuzzy controllers based on reinforcement learning for autonomous vehicle control," *IEEE Trans. Intell. Transp. Syst.*, vol. 6, no. 3, pp. 285–293, Sep. 2005.

[22] D. Zhao, Z. Hu, Z. Xia, C. Alippi, Y. Zhu, and D. Wang, "Full-range adaptive cruise control based on supervised adaptive dynamic programming," *Neurocomputing*, vol. 125, pp. 57–67, Feb. 2014.

[23] B. Dong, F. Zhou, K. Liu, and Y. Li, "Torque sensorless decentralized neuro-optimal control for modular and reconfigurable robots with uncertain environments," *Neurocomputing*, vol. 282, pp. 60–73, Mar. 2018.

[24] I. Grondman, L. Busoniu, G. A. D. Lopes, and R. Babuska, "A survey of actor-critic reinforcement learning: Standard and natural policy gradients," *IEEE Trans. Syst., Man, Cybern., C (Appl. Rev.)*, vol. 42, no. 6, pp. 1291–1307, Nov. 2012.

[25] D. Zhao, Y. Chen, and L. Lv, "Deep reinforcement learning with visual attention for vehicle classification," *IEEE Trans. Cognit. Develop. Syst.*, vol. 9, no. 4, pp. 356–367, Dec. 2017.

[26] B. Dong, F. Zhou, K. Liu, and Y. Li, "Decentralized robust optimal control for modular robot manipulators via critic-identifier structure-based adaptive dynamic programming," *Neural Comput. Appl.*, vol. 32, no. 8, pp. 3441–3458, Apr. 2020.

[27] D. Wang, H. He, and D. Liu, "Adaptive critic nonlinear robust control: A survey," *IEEE Trans. Cybern.*, vol. 47, no. 10, pp. 3429–3451, Oct. 2017.

[28] N. Wang and X.-J. Li, "Optimal output synchronization control of a class of complex dynamical networks with partially unknown system dynamics," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 51, no. 2, pp. 822–832, Feb. 2021.

[29] X.-J. Li and N. Wang, "Data-driven fault estimation and control for unknown discrete-time systems via multiobjective optimization method," *IEEE Trans. Cybern.*, early access, Aug. 12, 2020, doi: 10.1109/TCYB.2020.3010222.

[30] W. Bai, T. Li, and S. Tong, "NN reinforcement learning adaptive control for a class of nonstrict-feedback discrete-time systems," *IEEE Trans. Cybern.*, vol. 50, no. 11, pp. 4573–4584, Nov. 2020.

[31] H. Li, Y. Wu, and M. Chen, "Adaptive fault-tolerant tracking control for discrete-time multiagent systems via reinforcement learning algorithm," *IEEE Trans. Cybern.*, vol. 51, no. 3, pp. 1163–1174, Mar. 2021.

[32] Q. Yang and S. Jagannathan, "Reinforcement learning controller design for affine nonlinear discrete-time systems using online approximators," *IEEE Trans. Syst., Man, Cybern., B (Cybern.)*, vol. 42, no. 2, pp. 377–390, Apr. 2012.

[33] W. Bai, B. Zhang, Q. Zhou, and R. Lu, "Multigradient recursive reinforcement learning NN control for affine nonlinear systems with unmodeled dynamics," *Int. J. Robust Nonlinear Control*, vol. 30, no. 4, pp. 1643–1663, Mar. 2020.

[34] W. Bai, Q. Zhou, T. Li, and H. Li, "Adaptive reinforcement learning neural network control for uncertain nonlinear system with input saturation," *IEEE Trans. Cybern.*, vol. 50, no. 8, pp. 3433–3443, Aug. 2020.

[35] X. Zhong and H. He, "An event-triggered ADP control approach for continuous-time system with unknown internal states," *IEEE Trans. Cybern.*, vol. 47, no. 3, pp. 683–694, Mar. 2017.

[36] L. Xing, C. Wen, Z. Liu, H. Su, and J. Cai, "Event-triggered adaptive control for a class of uncertain nonlinear systems," *IEEE Trans. Autom. Control*, vol. 62, no. 4, pp. 2071–2076, Apr. 2017.

[37] X. Yang and H. He, "Adaptive critic designs for event-triggered robust control of nonlinear systems with unknown dynamics," *IEEE Trans. Cybern.*, vol. 49, no. 6, pp. 2255–2267, Jun. 2019.

[38] C. Mu and K. Wang, "Approximate-optimal control algorithm for constrained zero-sum differential games through event-triggering mechanism," *Nonlinear Dyn.*, vol. 95, no. 4, pp. 2639–2657, Mar. 2019.

[39] C. Mu, D. Wang, C. Sun, and Q. Zong, "Robust adaptive critic control design with network-based event-triggered formulation," *Nonlinear Dyn.*, vol. 90, no. 3, pp. 2023–2035, Nov. 2017.

[40] L. Cao, H. Li, N. Wang, and Q. Zhou, "Observer-based event-triggered adaptive decentralized fuzzy control for nonlinear large-scale systems," *IEEE Trans. Fuzzy Syst.*, vol. 27, no. 6, pp. 1201–1214, Jun. 2019.

[41] H. Liang, X. Guo, Y. Pan, and T. Huang, "Event-triggered fuzzy bipartite tracking control for network systems based on distributed reduced-order observers," *IEEE Trans. Fuzzy Syst.*, vol. 29, no. 6, pp. 1601–1614, Jun. 2021.

[42] M. Wang, Z. Wang, Y. Chen, and W. Sheng, "Event-based adaptive neural tracking control for discrete-time stochastic nonlinear systems: A triggering threshold compensation strategy," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 6, pp. 1968–1981, Jun. 2020.

[43] Y. Liu, Y. Ma, and Y. Wang, "Reliable sliding mode finite-time control for discrete-time singular Markovian jump systems with sensor fault and randomly occurring nonlinearities," *Int. J. Robust Nonlinear Control*, vol. 28, no. 2, pp. 381–402, Jan. 2018.

[44] L. Cao, H. Li, G. Dong, and R. Lu, "Event-triggered control for multiagent systems with sensor faults and input saturation," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 51, no. 6, pp. 3855–3866, Jun. 2021, doi: 10.1109/TSMC.2019.2938216.

[45] X.-J. Li and G.-H. Yang, "Adaptive decentralized control for a class of interconnected nonlinear systems via backstepping approach and graph theory," *Automatica*, vol. 76, pp. 87–95, Feb. 2017.

[46] B. Fan, Q. Yang, S. Jagannathan, and Y. Sun, "Output-constrained control of nonaffine multiagent systems with partially unknown control directions," *IEEE Trans. Autom. Control*, vol. 64, no. 9, pp. 3936–3942, Sep. 2019.

[47] S. S. Ge, J. Zhang, and T. H. Lee, "Adaptive neural network control for a class of MIMO nonlinear systems with disturbances in discrete-time," *IEEE Trans. Syst., Man Cybern., B (Cybern.)*, vol. 34, no. 4, pp. 1630–1645, Aug. 2004.

[48] Q. Zhou, S. Zhao, H. Li, R. Lu, and C. Wu, "Adaptive neural network tracking control for robotic manipulators with dead zone," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 12, pp. 3611–3620, Dec. 2019.

[49] G. Jiang, H. He, P. Xie, and Y. Tang, "Stacked multilevel-denoising autoencoders: A new representation learning approach for wind turbine gearbox fault diagnosis," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 9, pp. 2391–2402, Sep. 2017.

[50] H. Ma, H. Li, R. Lu, and T. Huang, "Adaptive event-triggered control for a class of nonlinear systems with periodic disturbances," *Sci. China Inf. Sci.*, vol. 63, no. 5, May 2020, Art. no. 150212.

[51] Y. Wu, Y. Pan, M. Chen, and H. Li, "Quantized adaptive finite-time bipartite NN tracking control for stochastic multiagent systems," *IEEE Trans. Cybern.*, vol. 51, no. 6, pp. 2870–2881, Jun. 2021.

[52] J. Li, Q. Yang, B. Fan, and Y. Sun, "Robust state/output-feedback control of coaxial-rotor MAVs based on adaptive NN approach," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 12, pp. 3547–3557, Dec. 2019.

[53] H. Ma, H. Ren, Q. Zhou, R. Lu, and H. Li, "Approximation-based Nussbaum gain adaptive control of nonlinear systems with periodic disturbances," *IEEE Trans. Syst., Man, Cybern., Syst.*, early access, Jan. 29, 2021, doi: 10.1109/TSMC.2021.3050993.

[54] Z. Ni, H. He, J. Wen, and X. Xu, "Goal representation heuristic dynamic programming on maze navigation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, no. 12, pp. 2038–2050, Dec. 2013.

[55] Y.-J. Liu, L. Tang, S. Tong, C. L. P. Chen, and D.-J. Li, "Reinforcement learning design-based adaptive tracking control with less learning parameters for nonlinear discrete-time MIMO systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 1, pp. 165–176, Jan. 2015.

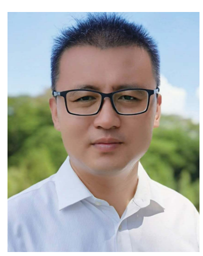
[56] C. Mu, Z. Ni, C. Sun, and H. He, "Data-driven tracking control with adaptive dynamic programming for a class of continuous-time nonlinear systems," *IEEE Trans. Cybern.*, vol. 47, no. 6, pp. 1460–1470, Jun. 2017.

[57] Z. Ni, N. Malla, and X. Zhong, "Prioritizing useful experience replay for heuristic dynamic programming-based learning systems," *IEEE Trans. Cybern.*, vol. 49, no. 11, pp. 3911–3922, Nov. 2019.

[58] C. Mu, Z. Ni, C. Sun, and H. He, "Air-breathing hypersonic vehicle tracking control based on adaptive dynamic programming," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 3, pp. 584–598, Mar. 2017.

[59] Y. Liu, X. Liu, Y. Jing, X. Chen, and J. Qiu, "Direct adaptive preassigned finite-time control with time-delay and quantized input using neural network," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 4, pp. 1222–1231, Apr. 2020.

[60] D.-J. Li and L. Tang, "Adaptive control for a class of chemical reactor systems in discrete-time form," *Neural Comput. Appl.*, vol. 24, nos. 7–8, pp. 1807–1814, Jun. 2014.



Hongyi Li (Senior Member, IEEE) received the Ph.D. degree in intelligent control from the University of Portsmouth, Portsmouth, U.K., in 2012.

He was a Research Associate with the Department of Mechanical Engineering, University of Hong Kong, Hong Kong, and Hong Kong Polytechnic University, Hong Kong. He was a Visiting Principal Fellow with the Faculty of Engineering and Information Sciences, University of Wollongong, Wollongong, NSW, Australia. He is currently a Professor with the Guangdong University of Technology, Guangdong, China. His research interests include intelligent control, cooperative control, sliding mode control, and their applications.

Dr. Li is a member of the IFAC Technical Committee on Computational Intelligence in Control. He was a recipient of the 2016 and 2019 Andrew P. Sage Best Transactions Paper Awards from the IEEE System, Man, Cybernetics Society, the Best Paper Award in Theory from ICCSS 2017, and the Zadeh Best Student Paper from IEEE ICCSS 2019, respectively. He has been in the Editorial Board of several international journals, including the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, the IEEE TRANSACTIONS ON FUZZY SYSTEMS, the IEEE TRANSACTIONS ON SYSTEMS, MAN AND CYBERNETICS: SYSTEMS, the IEEE TRANSACTIONS ON COGNITIVE AND DEVELOPMENTAL SYSTEMS, *SCIENCE CHINA Information Sciences*, the IEEE/CAA JOURNAL OF AUTOMATICA SINICA, *Journal of Systems Science and Complexity*, *Neural Networks*, *Asian Journal of Control, Circuits, Systems and Signal Processing*, and *International Journal of Control, Automation and Systems*. He has been Guest Editors of IEEE TRANSACTIONS ON CYBERNETICS and *IET Control Theory and Applications*.

Ying Wu received the B.S. degree in automation from Shandong Technology and Business University, Yantai, China, in 2018. She is currently pursuing the M.S. degree in detection technology and automatic equipment with Bohai University, Jinzhou, China.

Her current research interests include multiagent systems, adaptive fuzzy/neural control, and reinforcement learning.

Mou Chen (Senior Member, IEEE) received the B.Sc. degree in material science and engineering from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 1998, and the Ph.D. degree in automatic control engineering from the Nanjing University of Aeronautics and Astronautics, Nanjing, in 2004.

He was an Academic Visitor with the Department of Aeronautical and Automotive Engineering, Loughborough University, Loughborough, U.K., from November 2007 to February 2008. From June 2008 to September 2009, he was a Research Fellow with the Department of Electrical and Computer Engineering, National University of Singapore. He was a Senior Academic Visitor with the School of Electrical and Electronic Engineering, The University of Adelaide, Adelaide, SA, Australia, from May 2014 to November 2014. He is currently a Professor with the College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, China. His research interests include nonlinear system control, intelligent control, and flight control.

Renquan Lu (Senior Member, IEEE) received the Ph.D. degree in control science and engineering from Zhejiang University, Hangzhou, China, in 2004.

He is currently a Professor with the School of Automation, Guangdong University of Technology, Guangzhou, China. His research was supported by the National Science Fund for Distinguished Young Scientists of China in 2014, honored as the Distinguished Professor of Pearl River Scholars Program of Guangdong Province in 2015, and the Yangtze River Scholars Program by the Ministry of Education of China in 2017. His current research interests include complex systems, networked control systems, and nonlinear systems.