Distributed Robust Consensus Control of Nonlinear Multi-Agent Systems via Adaptive Dynamic Programming

Yongwei Zhang, Bo Zhao, *Senior Member, IEEE*, Derong Liu, *Fellow, IEEE*, and Shunchao Zhang

***Abstract*—This paper investigates the distributed robust consensus control (DRCC) of multi-agent systems via adaptive dynamic programming. To deal with matched uncertainties, a novel local value function is designed for each agent which contains the bound functions, the consensus errors, and the control laws of the follower and its neighbors. Subsequently, the distributed robust consensus control problem is transformed into an optimal consensus control problem. It shows that the distributed optimal consensus controller can achieve the robust consensus between the leader and the followers, which means that all followers subject to matched uncertainties synchronize to the leader. Hereafter, the distributed optimal consensus control law is derived by solving the coupled Hamilton-Jacobi-Bellman equation of each follower via critic-only structure. Furthermore, the convergence of the consensus error of all followers are guaranteed to be asymptotically stable by using Lyapunov's direct method. Finally, single link robot arms are adopted to verify the effectiveness of the proposed DRCC approach.**

***Index Terms*—Adaptive dynamic programming, multi-agent systems, robust control, optimal consensus control.**

I. Introduction

DISTRIBUTED control of multi-agent systems (MASs) has attracted much attention due to its broad applications in various areas, such as power systems, sensor networks, spacecraft systems and robotic systems. As one of the most important and fundamental issues in distributed control, consensus control which aims to make all the agents reach synchronization, is widely investigated in control community. In practice, each following agent not only has to be consistent with the leader, but also needs to consider the energy consumption, so it desires to achieve consensus in an optimal manner. Consequently, the distributed optimal consensus control (DOCC) receives extensively attention and aims to design distributed control laws such that all followers agree with the leader and minimize their performance index function. It is worth mentioning that DOCC needs to address the coupled Hamilton-Jacobi-Bellman (HJB) equation, which is a nonlinear partial differential equation and difficult to obtain its analytic solution [1]–[5].

This work was supported by the IEEE Computational Intelligence Society Graduate Student Research Grant 2021.

Yongwei Zhang, Derong Liu, and Shunchao Zhang are with the School of Automation, Guangdong University of Technology, Guangzhou 510006, China. emails: Yongwei_Zhang@mail2.gdut.edu.cn; derong@gdut.edu.cn; 1111904006@mail2.gdut.edu.cn.

Bo Zhao is with the School of Systems Science, Beijing Normal University, Beijing 100875, China. email: zhaobo@bnu.edu.cn.

It is well-known that adaptive dynamic programming (ADP), which was proposed by Werbos [6], has been recognized as an effective tool to solve the HJB equation and was adopted to address various control problems such as fault-tolerant [7], input constraints [8], dynamic uncertainties [9], and differential game [10]. For DOCC of MASs, several results have been developed. Zhang *et al.* [11] investigated the DOCC of discrete-time (DT) MASs via data-based reinforcement learning (RL) approach. Yang *et al.* [12] developed policy gradient-based RL method for DT MASs by using offline system interaction data. Sun *et al.* [13] studied the DOCC of continuous-time (CT) MASs subject to external disturbance. To save the computing resources, an event-triggered condition is designed for each agent and the controller is updated at triggering instants only. Wang *et al.* [14] proposed an off-policy model-free integral RL approach to address the fully DOCC problem of CT MASs. Khankalantary *et al.* [15] considered the DOCC of MASs with input saturation and collision avoidance constraints.

It is noticed that aforementioned results do not consider dynamic uncertainties. However, in practice, MASs works in complex environments, such as unmanned aerial vehicle in disaster relief, underwater robot in search and rescue, and robot arm in industrial production. Therefore, the occurrence of dynamic uncertainties is inevitable. In order to deal with the dynamic uncertainties, researchers have developed a large amount of ADP-based robust control methods in recent years. Jiang *et al.* [16] developed online robust ADP approach for nonlinear CT systems based on the small-gain theorem and the ADP theory. Liu *et al.* [17] studied the robust control of uncertain CT systems with input constraints. By designing a appropriate cost function, the robust control problem was transformed to an optimal control problem and a RL-based robust control approach was developed. Yang *et al.* [18] investigated the event-triggered robust control of unknown CT systems via adaptive critic designs. Wang *et al.* [19] proposed an event-based guaranteed cost control approach for CT systems with matched uncertainties. Mu *et al.* [20] addressed the robust tracking control problem with mismatched uncertainties by establishing an auxiliary system.

All the above mentioned works consider single agent systems only. However, MASs subject to dynamic uncertainties is more difficult to address since agents are connected with each other through a communication network, the effect of one agent's uncertainties will affect other agents. In conclusion, it is significant to develop an ADP-based DRCC scheme for MASs with dynamic uncertainties. This motivates our research.

The rest of this paper is organized as follows. In Section II, the DRCC problem of MASs is formulated. In Section III, the ADP-based DRCC design, the neural network implementation, and the stability analysis of MASs are provided. In Section IV, the effectiveness of the developed ADP-based DRCC method is verified on single link robot arms. In Section V, a brief conclusion is given.

II. Preliminaries

A. Graph Theory

Consider the undirected communication topology graph denoted by $\Pi_g = \{\mathcal{P}, \xi, \mathcal{A}\}$, where $\mathcal{P} = \{p_1, ..., p_N\}$ is a set of nodes, $\xi = \{(p_i, p_j)\colon p_i, p_j \in \mathcal{P}\}$ is a set of edges, $\mathcal{A} = [a_{ij}]$ is a weighted adjacency matrix, and N is the number of follower. If and only if the agent i and the agent j are directly connected, then $(p_i, p_j) \in \xi$. Moreover, $a_{ij} > 0$ if $(p_i, p_j) \in \xi$, $a_{ij} = 0$ if $(p_i, p_j) \notin \xi$, and $a_{ii} = 0$ for all $i = 1, ..., N$. Let $N_i = \{j\colon (p_i, p_j) \in \xi, j \neq i\}$ be a set of neighbors of agent i, $\bar{N}_i$ be a set of agent i and its neighbors, $\mathcal{D} = \mathrm{diag}\{d_1, ..., d_N\}$ with $d_i = \sum_{j\in N_i} a_{ij}$ be the degree matrix of Π_g, $\mathcal{L} = \mathcal{D} - \mathcal{A} = [l_{ij}]$ be the Laplacian matrix with $l_{ij} = -a_{ij}$ and $l_{ii} = \sum_{j=1}^{N_i} a_{ij}$.

B. Problem Formulation

Consider the nonlinear MASs with one leader and N followers. The system dynamics of each follower is described as

$$\dot{x}_i = \mathcal{F}_i(x_i) + \mathcal{G}_i(x_i)\big(u_i + \Pi_i(x_i)\big), \tag{1}$$

where $x_i \in \mathbb{R}^{n_i}$ is the system states of the follower i, $u_i \in \mathbb{R}^{m_i}$ is the control input of the follower i, $\Pi_i(x_i) \in \mathbb{R}^{m_i}$ is the matched uncertainty, and $\mathcal{F}_i(x_i) \in \mathbb{R}^{n_i}$ and $\mathcal{G}_i(x_i) \in \mathbb{R}^{n_i\times m_i}$ are nonlinear system functions.

Assumption 1: The matched uncertainty $\Pi_i(x_i)$ is norm-bounded as $\|\Pi_i(x_i)\| \leq \bar{\Pi}_i(x_i)$, where $\bar{\Pi}_i(x_i)$ is a positive definite function.

Assumption 2: The control input matrix $\mathcal{G}_i(x_i)$ is norm-bounded as $\|\mathcal{G}_i(x_i)\| \leq \bar{\mathcal{G}}_i$, where $\bar{\mathcal{G}}_i$ is a positive constant.

The system dynamics of the leader is given by

$$\dot{x}_0 = f_0(x_0), \tag{2}$$

where $x_0 \in \mathbb{R}^{n_0}$ and $f_0(x_0) \in \mathbb{R}^{n_0}$ is a differential function. Then, the local neighborhood consensus error of the follower i is defined as

$$\mathcal{E}_i = \sum_{j\in N_i} a_{ij}(x_i - x_j) + c_i(x_i - x_0), \tag{3}$$

where $c_i \geq 0$ is the pinning gain. The dynamics of the local neighborhood consensus error can be obtained by differentiating (3) as

$$\begin{aligned}
\dot{\mathcal{E}}_i =& \sum_{j\in N_i} a_{ij}(\dot{x}_i - \dot{x}_j) + c_i(\dot{x}_i - \dot{x}_0)\\
=& \sum_{j\in N_i} a_{ij}\Big(\mathcal{F}_i(x_i) + \mathcal{G}_i(x_i)\big(u_i + \Pi(x_i)\big)\\
&- \mathcal{F}_j(x_j) - \mathcal{G}_j(x_j)\big(u_j + \Pi(x_j)\big)\Big)\\
&+ c_i\Big(\mathcal{F}_i(x_i) + \mathcal{G}_i(x_i)\big(u_i + \Pi(x_i)\big) - f_0(x_0)\Big)\\
=& \sum_{j\in N_i} a_{ij}\Big(\mathcal{F}_i(x_i) + \mathcal{G}_i(x_i)\big(u_i + \Pi(x_i)\big)\Big)\\
&+ c_i\Big(\mathcal{F}_i(x_i) + \mathcal{G}_i(x_i)\big(u_i + \Pi(x_i)\big)\Big)\\
&- \sum_{j\in N_i} a_{ij}\Big(\mathcal{F}_j(x_j) + \mathcal{G}_j(x_j)\big(u_j + \Pi(x_j)\big)\Big)\\
&- c_i f_0(x_0)\\
=& (l_{ii} + c_i)\Big(\mathcal{F}_i(x_i) + \mathcal{G}_i(x_i)\big(u_i + \Pi(x_i)\big)\Big)\\
&- \sum_{j\in N_i} a_{ij}\Big(\mathcal{F}_j(x_j) + \mathcal{G}_j(x_j)\big(u_j + \Pi(x_j)\big)\Big)\\
&- c_i f_0(x_0).
\end{aligned} \tag{4}$$

The nominal system of (1) is expressed by

$$\dot{x}_i = \mathcal{F}_i(x_i) + \mathcal{G}_i(x_i)u_i. \tag{5}$$

Then, the corresponding local neighborhood consensus error dynamics is given by

$$\begin{aligned}
\dot{\mathcal{E}}_i =& (l_{ii} + c_i)\big(\mathcal{F}_i(x_i) + \mathcal{G}_i(x_i)u_i\big) - c_i f_0(x_0)\\
&- \sum_{j\in N_i} a_{ij}\big(\mathcal{F}_j(x_j) + \mathcal{G}_j(x_j)u_j\big).
\end{aligned} \tag{6}$$

The local value function of the follower i is defined as

$$\mathcal{V}_i(\mathcal{E}_i) = \int_t^\infty \Big(\Gamma_i(x_i) + \mathcal{C}_i\big(\mathcal{E}_i(\tau), u_i(\tau), u_{(-i)}(\tau)\big)\Big)d\tau,$$

where $\Gamma_i(x_i)$ is a positive definite upper bound function which will be designed in next section and the utility function $\mathcal{C}_i(\cdot)$ is designed as

$$\mathcal{C}_i(\mathcal{E}_i, u_i, u_{(-i)}) = \mathcal{E}_i^{\mathsf{T}} Q_{ii}\mathcal{E}_i + u_i^{\mathsf{T}} R_{ii} u_i + \sum_{j\in N_i} u_j^{\mathsf{T}} R_{ij} u_j, \tag{7}$$

where $u_{(-i)} = \{u_j | j \in N_i\}$ are the control inputs of the neighbors of follower i, $Q_{ii} \in \mathbb{R}^{n_i\times n_i}$, $R_{ii} \in \mathbb{R}^{m_i\times m_i}$ and $R_{ij} \in \mathbb{R}^{m_j\times m_j}$ are positive definite matrices. The Hamiltonian of the follower i is defined as

$$\begin{aligned}
&\mathcal{H}_i(\mathcal{E}_i, \nabla\mathcal{V}_i(\mathcal{E}_i), u_i, u_{(-i)})\\
&= \Gamma_i(x_i) + \mathcal{C}_i(\mathcal{E}_i, u_i, u_{(-i)})\\
&\quad + \nabla\mathcal{V}_i^{\mathsf{T}}(\mathcal{E}_i)\Bigg((l_{ii} + c_i)\big(\mathcal{F}_i(x_i) + \mathcal{G}_i(x_i)u_i\big)\\
&\quad - c_i f_0(x_0) - \sum_{j\in N_i} a_{ij}\big(\mathcal{F}_j(x_j) + \mathcal{G}_j(x_j)u_j\big)\Bigg).
\end{aligned}$$

The local optimal value function of the follower i

$$\begin{aligned}
&\mathcal{V}_i^*(\mathcal{E}_i)\\
&= \min_{u_i\in\Re(\Omega)} \int_t^\infty \Big(\Gamma_i(x_i) + \mathcal{C}_i\big(\mathcal{E}_i(\tau), u_i(\tau), u_{(-i)}(\tau)\big)\Big)d\tau
\end{aligned}$$

satisfies the HJB equation as

$$\min_{u_i\in\Re(\Omega)} \mathcal{H}_i\big(\mathcal{E}_i, \nabla\mathcal{V}_i^*(\mathcal{E}_i), u_i, u_{(-i)}\big) = 0, \tag{8}$$

where $\Re(\Omega)$ is a set of admissible control. Then, the local optimal consensus control law is derived by

$$u_i^* = -\frac{d_i + c_i}{2} R_{ii}^{-1} \mathcal{G}_i^{\mathsf{T}}(x_i) \nabla \mathcal{V}_i^*(\mathcal{E}_i). \tag{9}$$

Based on (8) and (9), we can obtain

$$\begin{aligned} 0 = {}& \Gamma_i(x_i) + \mathcal{C}_i\big(\mathcal{E}_i, u_i^*, u_{(-i)}^*\big) \\ &+ \nabla \mathcal{V}_i^{*\mathsf{T}}(\mathcal{E}_i)\bigg((l_{ii} + c_i)\big(\mathcal{F}_i(x_i) + \mathcal{G}_i(x_i)u_i^*\big) \\ &- c_i f_0(x_0) - \sum_{j\in N_i} a_{ij}\big(\mathcal{F}_j(x_j) + \mathcal{G}_j(x_j)u_j^*\big)\bigg). \end{aligned} \tag{10}$$

Noticing that (10) is the coupled HJB equation, which is difficult to solve due to its high nonlinearities [22]. In the next section, an ADP-based DRCC method is proposed to overcome this bottleneck.

III. Distributed Robust Consensus Control Design

A. Distributed Robust Consensus Control Approach

In this section, the positive definite upper bound function $\Gamma_i(x_i)$ is deigned as

$$\begin{aligned} \Gamma_i(x_i) = {}& \frac{1}{2}\bar{\mathcal{G}}_i^2 (l_{ii} + c_i)^2 \bar{\Pi}_i^2(x_i) \\ &+ \frac{1}{2} N_i a_{ij}^2 \bar{\mathcal{G}}_i^2 \sum_{j\in N_i} \bar{\Pi}_j^2(x_j). \end{aligned}$$

Then, the DRCC of MASs with matched uncertainties is transformed to DOCC of its nominal form. The equivalence of this problem transformation is analyzed in the following theorem.

Theorem 1: Consider the nonlinear MASs with the leader (2) and followers (1), the dynamics of the local neighborhood consensus error given by (4), the local optimal consensus control law given by (9), and the Assumptions 1 and 2, if there exists a matrix Q_{ii} satisfying

$$\varrho^2 \lambda_{\min}(Q_{ii}) \|\mathcal{E}_i\|^2 \geq \left\|\nabla \mathcal{V}_i^{\mathsf{T}}(\mathcal{E}_i)\right\|^2, \tag{11}$$

where $0 < \varrho < 1$. Then, the local neighborhood consensus error of each follower is asymptotically stable.

Proof. Select the Lyapunov function candidate as

$$\mathcal{L}_{T1} = \mathcal{V}_i(\mathcal{E}_i). \tag{12}$$

Differentiating $\mathcal{L}_{T1}$ along the solution of (4), we can get

$$\begin{aligned} \dot{\mathcal{L}}_{T1} = {}& \nabla \mathcal{V}_i^{\mathsf{T}}(\mathcal{E}_i)\bigg((l_{ii} + c_i)\Big(\mathcal{F}_i(x_i) + \mathcal{G}_i(x_i)\big(u_i^* + \Pi_i(x_i)\big)\Big) \\ &- \sum_{j\in N_i} a_{ij}\Big(\mathcal{F}_j(x_j) + \mathcal{G}_j(x_j)\big(u_j^* + \Pi_j(x_j)\big)\Big) \\ &- c_i f_0(x_0)\bigg) \end{aligned} \tag{13}$$

Based on (10), we have

$$\begin{aligned} &-\Gamma_i(x_i) - \mathcal{C}_i\big(\mathcal{E}_i, u_i^*, u_{(-i)}^*\big) \\ &= \nabla \mathcal{V}_i^{\mathsf{T}}(\mathcal{E}_i)\bigg((l_{ii} + c_i)\Big(\mathcal{F}_i(x_i) + \mathcal{G}_i(x_i)u_i^*\Big) \\ &- \sum_{j\in N_i} a_{ij}\Big(\mathcal{F}_j(x_j) + \mathcal{G}_j(x_j)u_j^*\Big) - c_i f_0(x_0)\bigg). \end{aligned} \tag{14}$$

Substituting (14) into (13), $\dot{\mathcal{L}}_{T1}$ becomes

$$\begin{aligned} \dot{\mathcal{L}}_{T1} = {}& -\Gamma_i(x_i) - \mathcal{C}_i\big(\mathcal{E}_i, u_i, u_{(-i)}\big) \\ &+ \nabla \mathcal{V}_i^{\mathsf{T}}(\mathcal{E}_i)(l_{ii} + c_i)\mathcal{G}_i(x_i)\Pi_i(x_i) \\ &- \nabla \mathcal{V}_i^{\mathsf{T}}(\mathcal{E}_i)\sum_{j\in N_i} a_{ij}\mathcal{G}_j(x_j)\Pi_j(x_j) \\ \leq {}& -\Gamma_i(x_i) - \mathcal{E}_i^{\mathsf{T}} Q_{ii}\mathcal{E}_i + \left\|\nabla \mathcal{V}_i^{\mathsf{T}}(\mathcal{E}_i)\right\|^2 \\ &+ \frac{1}{2}\left\|(l_{ii} + c_i)\mathcal{G}_i(x_i)\Pi_i(x_i)\right\|^2 \\ &+ \frac{1}{2}\left\|\sum_{j\in N_i} a_{ij}\mathcal{G}_j(x_j)\Pi_j(x_j)\right\|^2 \\ \leq {}& -\varrho^2\lambda_{\min}(Q_{ii})\|\mathcal{E}_i\|^2 + (\varrho^2 - 1)\lambda_{\min}(Q_{ii})\|\mathcal{E}_i\|^2 \\ &+ \frac{1}{2}\bar{\mathcal{G}}_i^2(l_{ii} + c_i)^2\bar{\Pi}_i^2(x_i) + \left\|\nabla \mathcal{V}_i^{\mathsf{T}}(\mathcal{E}_i)\right\|^2 - \Gamma_i(x_i) \\ &+ \frac{1}{2}N_i a_{ij}^2\bar{\mathcal{G}}_i^2\sum_{j\in N_i}\bar{\Pi}_j^2(x_j) \\ \leq {}& -\varrho^2\lambda_{\min}(Q_{ii})\|\mathcal{E}_i\|^2 + (\varrho^2 - 1)\lambda_{\min}(Q_{ii})\|\mathcal{E}_i\|^2 \\ &+ \left\|\nabla \mathcal{V}_i^{\mathsf{T}}(\mathcal{E}_i)\right\|^2. \end{aligned} \tag{15}$$

Therefore, $\dot{\mathcal{L}}_{T1} < 0$ if the condition (11) holds. It means that the local neighborhood consensus error of each follower is guaranteed to be asymptotically stable. The proof is completed.

B. Neural Network Implementation

In this section, critic NNs are adopted to obtain the approximate solution of coupled HJB equations. According to the universal approximation property of NN, the local optimal value function of the follower i is expressed as

$$\mathcal{V}_i^*(\mathcal{E}_i) = W_{ic}^{*\mathsf{T}}\sigma_{ic}(\mathcal{E}_i) + \varepsilon_{ic}(\mathcal{E}_i), \tag{16}$$

where $W_{ic}^* \in \mathbb{R}^{h_c}$ is the ideal weight vector, $\sigma_{ic}(\mathcal{E}_i) \in \mathbb{R}^{h_c}$ is the activation function, h_c is the number of hidden neurons, and $\varepsilon_{ic}(\mathcal{E}_i) \in \mathbb{R}$ is the reconstruction error. Then, the partial derivative of $\mathcal{V}_i^*(\mathcal{E}_i)$ with respect to $\mathcal{E}_i$ is given by

$$\nabla \mathcal{V}_i^*(\mathcal{E}_i) = \nabla\sigma_c^{\mathsf{T}}(\mathcal{E}_i)W_{ic}^* + \nabla\varepsilon_{ic}^{\mathsf{T}}(\mathcal{E}_i). \tag{17}$$

The approximate local value function is defined as

$$\hat{\mathcal{V}}_i(\mathcal{E}_i) = \hat{W}_{ic}^{\mathsf{T}}\sigma_{ic}(\mathcal{E}_i), \tag{18}$$

where $\hat{W}_{ic} \in \mathbb{R}^{h_c}$ is the estimate of W_{ic}. Similarly, we have

$$\nabla\hat{\mathcal{V}}_i(\mathcal{E}_i) = \nabla\sigma_{ic}^{\mathsf{T}}(\mathcal{E}_i)\hat{W}_{ic}. \tag{19}$$

According to (9) and (17), the local optimal consensus control law is expressed as

$$u_i^* = -\frac{d_i + c_i}{2}R_{ii}^{-1}\mathcal{G}_i^{\mathsf{T}}(x_i)\left(\nabla\sigma_{ic}^{\mathsf{T}}(\mathcal{E}_i)W_{ic}^* + \nabla\varepsilon_{ic}^{\mathsf{T}}(\mathcal{E}_i)\right). \tag{20}$$

Then, the approximate local consensus control law is given by

$$\hat{u}_i = -\frac{d_i + c_i}{2}R_{ii}^{-1}\mathcal{G}_i^{\mathsf{T}}(x_i)\nabla\sigma_{ic}^{\mathsf{T}}(\mathcal{E}_i)\hat{W}_{ic}. \tag{21}$$

Based on (10) and (21), the approximate Hamiltonian is

$$\begin{aligned}&\hat{\mathcal{H}}_i(\mathcal{E}_i,\hat{W}_{ic},\hat{u}_i,\hat{u}_{(-i)})\\&\quad=\mathcal{E}_i^{\mathsf{T}}Q_{ii}\mathcal{E}_i+\hat{u}_i^{\mathsf{T}}R_{ii}\hat{u}_i+\sum_{j\in N_i}\hat{u}_j^{\mathsf{T}}R_{ij}\hat{u}_j\\&\qquad+\hat{W}_{ic}^{\mathsf{T}}\nabla\sigma_{ic}(\mathcal{E}_i)\Big((l_{ii}+c_i)\big(\mathcal{F}_i(x_i)+\mathcal{G}_i(x_i)\hat{u}_i\big)\\&\qquad-c_if_0(x_0)-\sum_{j\in N_i}a_{ij}\big(\mathcal{F}_j(x_j)+\mathcal{G}_j(x_j)\hat{u}_j\big)\Big)\\&\quad\triangleq e_{ic}.\end{aligned}\tag{22}$$

Let $\tilde{W}_{ic}=W_{ic}-\hat{W}_{ic}$ be the weight estimation error. The gradient descent algorithm is employed to minimize the target function $E_{ic}=\frac{1}{2}e_{ic}^{\mathsf{T}}e_{ic}$. Hence, the critic NN weight updating rule is given by

$$\dot{\hat{W}}_{ic}=-\alpha_c\frac{1}{(1+\Phi_i^{\mathsf{T}}\Phi_i)^2}\left(\frac{\partial E_{ic}}{\partial\hat{W}_{ic}}\right)=-\frac{\alpha_ce_{ic}\Phi_i}{(1+\Phi_i^{\mathsf{T}}\Phi_i)^2},\tag{23}$$

where $\alpha_c>0$ is the learning rate and $\Phi_i=\nabla\sigma_{ic}(\mathcal{E}_i)((l_{ii}+c_i)(\mathcal{F}_i(x_i)+\mathcal{G}_i(x_i)\hat{u}_i)-c_if_0(x_0)-\sum_{j\in N_i}a_{ij}(\mathcal{F}_j(x_j)+\mathcal{G}_j(x_j)\hat{u}_j))$

Theorem 2: Consider the nonlinear MASs with the leader (2) and followers (5), the dynamics of the local neighborhood consensus error given by (6), if the critic NN weight is updated by (23), then the weight approximation error $\tilde{W}_{ic}$ can be guaranteed to be UUB.

Proof. The proof of Theorem 2 has been provided in [21], [27], so it omitted here.

C. Stability Analysis

In this section, we will prove that the approximate local consensus control law (21) can guarantee the local neighborhood consensus error of each follower with matched uncertainties to be UUB. Before stability analysis, the following assumption which is common in ADP literature [23]–[26] is provided.

Assumption 3: $\nabla\sigma_{ic}(\mathcal{E}_i)$, $\nabla\varepsilon_{ic}(\mathcal{E}_i)$, $\tilde{W}_{ic}$ and W_{ic}^* are norm-bounded, i.e.,

$$\begin{aligned}&\|\nabla\sigma_{ic}(\mathcal{E}_i)\|\le\bar{\sigma}_{ic},\|\nabla\varepsilon_{ic}(\mathcal{E}_i)\|\le\bar{\varepsilon}_{ic},\\&\|\tilde{W}_{ic}\|\le\bar{W}_{ic},\|W_{ic}^*\|\le\bar{W}_{icM},\end{aligned}$$

where $\bar{\sigma}_{ic}$, $\bar{\varepsilon}_{ic}$, $\bar{W}_{ic}$ and $\bar{W}_{ic}^*$ are positive constants.

Theorem 3: Consider the nonlinear MASs with the leader (2) and followers (5), the dynamics of the local neighborhood consensus error given by (6), the critic NN weight updated by (23), and Assumptions 1–3. Then, the approximate local consensus control law (21) can guarantee the local neighborhood consensus error of each follower to be UUB.

Proof. Select the Lyapunov function candidate as

$$\mathcal{L}_{T2}=\mathcal{V}_i^*(\mathcal{E}_i).\tag{24}$$

Differentiating $\mathcal{L}_{T2}$ along the solution of (6), we can get

$$\begin{aligned}\dot{\mathcal{L}}_{T2}&=\nabla\mathcal{V}_i^{*\mathsf{T}}(\mathcal{E}_i)\dot{\mathcal{E}}_i\\&=\nabla\mathcal{V}_i^{*\mathsf{T}}(\mathcal{E}_i)\Bigg((l_{ii}+c_i)\big(\mathcal{F}_i(x_i)+\mathcal{G}_i(x_i)\hat{u}_i\big)-c_if_0(x_0)\\&\qquad-\sum_{j\in N_i}a_{ij}\big(\mathcal{F}_j(x_j)+\mathcal{G}_j(x_j)\hat{u}_j\big)\Bigg)\\&=\nabla\mathcal{V}_i^{*\mathsf{T}}(\mathcal{E}_i)(l_{ii}+c_i)\mathcal{G}_i(x_i)\hat{u}_i\\&\qquad-\nabla\mathcal{V}_i^{*\mathsf{T}}(\mathcal{E}_i)\sum_{j\in N_i}a_{ij}\mathcal{G}_j(x_j)\hat{u}_j\\&\qquad-\nabla\mathcal{V}_i^{*\mathsf{T}}(\mathcal{E}_i)(l_{ii}+c_i)\mathcal{G}_i(x_i)u_i^*\\&\qquad+\nabla\mathcal{V}_i^{*\mathsf{T}}(\mathcal{E}_i)\sum_{j\in N_i}a_{ij}\mathcal{G}_j(x_j)u_j^*\\&\qquad-\Gamma_i(x_i)-\mathcal{C}_i(\mathcal{E}_i,u_i^*,u_{(-i)}^*)\\&=\nabla\mathcal{V}_i^{*\mathsf{T}}(\mathcal{E}_i)(l_{ii}+c_i)\mathcal{G}_i(x_i)(\hat{u}_i-u_i^*)\\&\qquad+\nabla\mathcal{V}_i^{*\mathsf{T}}(\mathcal{E}_i)\sum_{j\in N_i}a_{ij}\mathcal{G}_j(x_j)(u_j^*-\hat{u}_j)\\&\qquad-\Gamma_i(x_i)-\mathcal{C}_i(\mathcal{E}_i,u_i^*,u_{(-i)}^*).\end{aligned}\tag{25}$$

According to (9), we can further drive that

$$\begin{aligned}\dot{\mathcal{L}}_{T2}\le&-2u_i^{*\mathsf{T}}R_{ii}(\hat{u}_i-u_i^*)+\frac{1}{2}\nabla\mathcal{V}_i^{*\mathsf{T}}(\mathcal{E}_i)\nabla\mathcal{V}_i^*(\mathcal{E}_i)\\&+\frac{1}{2}\Big(\sum_{j\in N_i}a_{ij}\mathcal{G}_j(x_j)(u_j^*-\hat{u}_j)\Big)^{\mathsf{T}}\\&\quad\times\Big(\sum_{j\in N_i}a_{ij}\mathcal{G}_j(x_j)(u_j^*-\hat{u}_j)\Big)\\&-\Gamma_i(x_i)-\mathcal{C}_i(\mathcal{E}_i,u_i^*,u_{(-i)}^*)\\\le&-2u_i^{*\mathsf{T}}R_{ii}\hat{u}_i+2u_i^{*\mathsf{T}}R_{ii}u_i^*+\frac{1}{2}\|\nabla\mathcal{V}_i^*(\mathcal{E}_i)\|^2\\&+\frac{1}{2}\left\|\sum_{j\in N_i}a_{ij}\mathcal{G}_j(x_j)(u_j^*-\hat{u}_j)\right\|^2\\&-\Gamma_i(x_i)-\mathcal{C}_i(\mathcal{E}_i,u_i^*,u_{(-i)}^*)\\\le&[u_i^*-\hat{u}_i]^{\mathsf{T}}R_{ii}[u_i^*-\hat{u}_i]-\hat{u}_i^{\mathsf{T}}R_{ii}\hat{u}_i\\&+\frac{1}{2}\|\nabla\sigma_{ic}^{\mathsf{T}}(\mathcal{E}_i)W_{ic}^*+\nabla\varepsilon_{ic}^{\mathsf{T}}(\mathcal{E}_i)\|^2\\&+\frac{N_i}{2}\sum_{j\in N_i}\|a_{ij}\mathcal{G}_j(x_j)(u_j^*-\hat{u}_j)\|^2-\mathcal{E}_i^{\mathsf{T}}Q_i\mathcal{E}_i\\\le&\lambda_{\max}(R_{ii})\|u_i^*-\hat{u}_i\|^2+\bar{\sigma}_{ic}^2\bar{W}_{icM}^2+\bar{\varepsilon}_{ic}^2\\&+\frac{N_i}{2}\sum_{j\in N_i}a_{ij}^2\bar{\mathcal{G}}_j^2\|u_j^*-\hat{u}_j\|^2-\mathcal{E}_i^{\mathsf{T}}Q_i\mathcal{E}_i.\end{aligned}\tag{26}$$

Noticing that

$$\begin{aligned}&\|\hat{u}_j-u_j^*\|^2\\&=\bigg\|-\frac{d_j+c_j}{2}R_{jj}^{-1}\mathcal{G}_j^{\mathsf{T}}(x_j)\nabla\sigma_{jc}^{\mathsf{T}}(\mathcal{E}_j)\hat{W}_{jc}\\&\quad+\frac{d_j+c_j}{2}R_{jj}^{-1}\mathcal{G}_j^{\mathsf{T}}(x_j)\nabla\sigma_{jc}^{\mathsf{T}}(\mathcal{E}_j)W_{jc}\\&\quad+\frac{d_j+c_j}{2}R_{jj}^{-1}\mathcal{G}_j^{\mathsf{T}}(x_j)\nabla\varepsilon_{jc}^{\mathsf{T}}(\mathcal{E}_j)\bigg\|^2\\&=\bigg\|\frac{d_j+c_j}{2}R_{jj}^{-1}\mathcal{G}_j^{\mathsf{T}}(x_j)\nabla\sigma_{jc}^{\mathsf{T}}(\mathcal{E}_j)\tilde{W}_{jc}\\&\quad+\frac{d_j+c_j}{2}R_{jj}^{-1}\mathcal{G}_j^{\mathsf{T}}(x_j)\nabla\varepsilon_{jc}^{\mathsf{T}}(\mathcal{E}_j)\bigg\|^2\\&\le\Big\|(d_j+c_j)R_{jj}^{-1}\mathcal{G}_j^{\mathsf{T}}(x_j)\nabla\sigma_{jc}^{\mathsf{T}}(\mathcal{E}_j)\tilde{W}_{jc}\Big\|^2\\&\quad+\Big\|(d_j+c_j)R_{jj}^{-1}\mathcal{G}_j^{\mathsf{T}}(x_j)\nabla\varepsilon_{jc}^{\mathsf{T}}(\mathcal{E}_j)\Big\|^2\end{aligned}$$

$$\leq (d_j + c_j)^2 \bar{R}_{jj}^2 \bar{\mathcal{G}}_j^2 \bar{\sigma}_{jc}^2 \bar{W}_{jc}^2 + (d_j + c_j)^2 \bar{R}_{jj}^2 \bar{\mathcal{G}}_j^2 \bar{\varepsilon}_{jc}^2. \quad (27)$$

Let $\Theta_j = (d_j+c_j)^2 \bar{R}_{jj}^2 \bar{\mathcal{G}}_j^2 \bar{\sigma}_{jc}^2 \bar{W}_{jc}^2 + (d_j+c_j)^2 \bar{R}_{jj}^2 \bar{\mathcal{G}}_j^2 \bar{\varepsilon}_{jc}^2$. Then, we further have

$$\begin{aligned} \dot{\mathcal{L}}_{T2} \leq & \lambda_{\max}(R_{ii})\Theta_i^2 + \bar{\sigma}_{ic}^2 \bar{W}_{icM}^2 + \bar{\varepsilon}_{ic}^2 \\ & + \frac{N_i}{2} \sum_{j \in N_i} a_{ij}^2 \bar{\mathcal{G}}_j^2 \Theta_j^2 - \varrho_1^2 \lambda_{\min}(Q_{ii}) \|\mathcal{E}_i\|^2 \\ & + (\varrho_1^2 - 1)\lambda_{\min}(Q_{ii}) \|\mathcal{E}_i\|^2, \end{aligned} \quad (28)$$

where $0 < \varrho_1 < 1$. Therefore, $\dot{\mathcal{L}}_{T2} < 0$ if $\mathcal{E}_i$ lies outside the compact set

$$\Omega_{\mathcal{E}_i} = \left\{ \mathcal{E}_i \colon \|\mathcal{E}_i\| \leq \sqrt{\frac{\lambda_2}{(1 - \varrho_1^2)\lambda_{\min}(Q_{ii})}} \right\}, \quad (29)$$

where $\lambda_2 = \lambda_{\max}(R_{ii})\Theta_i^2 + \bar{\sigma}_{ic}^2 \bar{W}_{icM}^2 + \bar{\varepsilon}_{ic}^2 + \frac{N_i}{2} \sum_{j \in N_i} a_{ij}^2 \bar{\mathcal{G}}_j^2 \Theta_j^2$. It means that the approximate local consensus control law (21) guarantees the UUB of the local neighborhood consensus error of each follower. The proof is completed.

IV. Numerical Simulation

In this section, single link robot arms are adopted to verify the effectiveness of the developed ADP-based DRCC scheme. Assume that MAS contains one leader and three followers. The communication topology of the MAS is shown in Fig. 1 and the corresponding parameter values are given as $c_1 = 1, c_2 = 0, c_3 = 0, a_{12} = 0.1, a_{13} = 0.1, a_{21} = 0.5, a_{23} = 0.5, a_{31} = 0.4$ and $a_{32} = 0.4$.

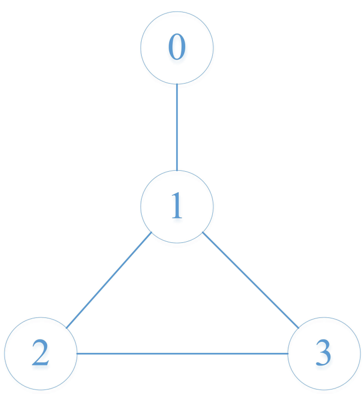

Fig. 1: Communication topology.

The dynamics of the single link robot arm i is given by

$$G\ddot{q}_i(t) - Mgh\sin\big(q_i(t)\big) + D\dot{q}_i(t) = u_i + d_i, \quad (30)$$

where q_i is the position of the joint, G is the moment of inertia, M is the quality of the connecting rod, h is the length of the arm, g is the acceleration of gravity, D is the viscous friction, u_i is the control input, and d_i is the matched uncertainty. The values of these parameters are provided in Table I.

TABLE I: Parameters of the single link robot arm

Parameter	G	M	h	g	D
Value	10	10	0.5	9.8	2

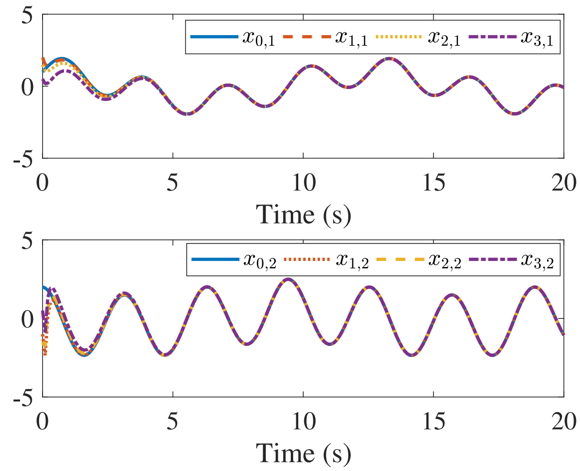

Fig. 2: Trajectories tracking.

Let $x_i = [x_{i,1}, x_{i,2}]^\mathsf{T} = [q_i, \dot{q}_i]^\mathsf{T} (i = 1, 2, 3)$. Then, the dynamics of the single link robot arm i is rewritten as

$$\begin{cases} \dot{x}_{i,1} = x_{i,2} \\ \dot{x}_{i,2} = -4.905\sin(x_{i,1}) - 0.2x_{i,2} + 0.1u_i + 0.1d_i \end{cases}$$

The trajectory of the leader is chosen as

$$x_0 = \begin{bmatrix} x_{0,1} \\ x_{0,2} \end{bmatrix} = \begin{bmatrix} \sin(2t) - \cos(0.5t) \\ 2\cos(2t) - 0.5\sin(0.5t) \end{bmatrix}. \quad (31)$$

We aim to make the joint position and the speed of all followers catch up with the leader's. Let $Q_{ii} = 10I_4$, $R_{ii} = 0.01I_1$, $R_{ij} = 0.05I_1$, the activation function of the critic NN be $\sigma_{ic} = [\mathcal{E}_{i1}^2, \mathcal{E}_{i2}^2, \mathcal{E}_{i1}\mathcal{E}_{i2}]$, the learning rate of the critic NN be $\alpha_c = 1$, the matched uncertainty be $d = 5\sin(x_{i,1})\cos(x_{i,2})$, and the upper bound function be $\bar{\Pi}(x_i) = \|x_i\|$.

Simulation results are given in Figs. 2–6. In Fig. 2, we can observe that the followers can track the leader within 10sec. Fig. 3 shows the consensus errors of all followers. It is clear that the consensus errors converge to small region of zero after 10sec. Fig. 4 displays that critic weight vectors will converge to $\hat{W}_{1c} = [27.32, 5.73, 67.09]^\mathsf{T}$, $\hat{W}_{2c} = [48.94, 48.54, 55.55]^\mathsf{T}$ and $\hat{W}_{3c} = [16.40, 47.03, 64.21]^\mathsf{T}$, respectively. Fig. 5 provides the trajectories of control inputs of all followers. Fig. 6 reveals that the developed ADP-based DRCC scheme can achieve the robust consensus of multi-single link robot arms.

V. Conclusion

In this paper, the ADP-based DRCC scheme is developed for MASs with matched uncertainties. To begin with, the DRCC problem is converted to a DOCC problem by designing a new value function for each follower. Theoretical analysis shows that distributed optimal consensus controllers can guarantee robust consensus of MASs with matched uncertainties. Meanwhile, the critic-only structure is established to obtain the approximate solution of the HJB equation of each follower. Moreover, the UUB stability of MASs with matched uncertainties is demonstrated by using Lyapunov's direct method. In the final, single link robot arms are employed to verify the effectiveness of the developed ADP-based DRCC scheme.

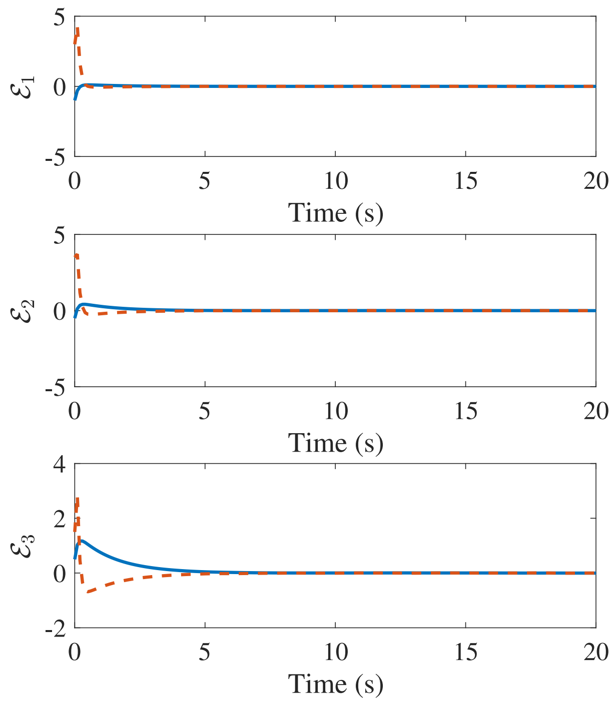

Fig. 3: Consensus errors.

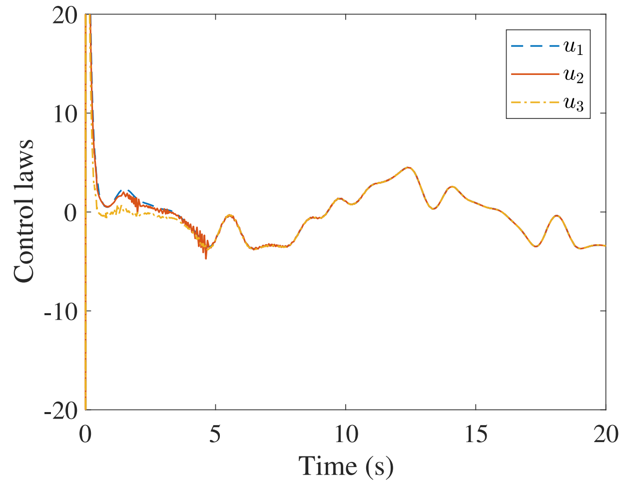

Fig. 5: Control inputs.

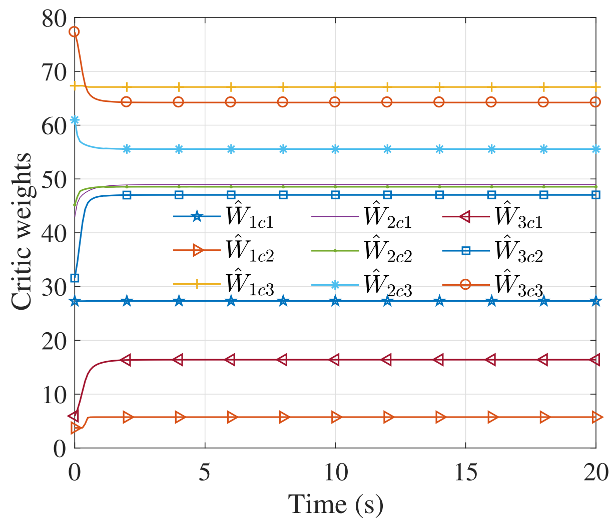

Fig. 4: Critic weights.

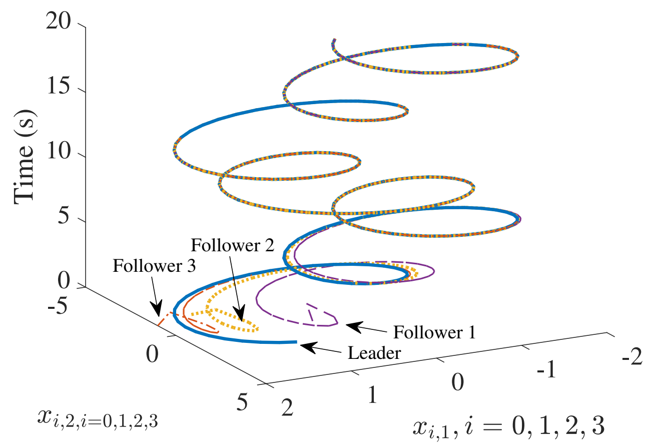

Fig. 6: Trajectories tracking with matched uncertainties.

REFERENCES

[1] D. Liu, Q. Wei, D. Wang, X. Yang, and H. Li, *Adaptive Dynamic Programming With Applications in Optimal Control*, Cham, Switzerland: Springer, 2017.

[2] Q. Wei, D. Liu, Y. Liu, and R. Song. "Optimal constrained self-learning battery sequential management in microgrid via adaptive dynamic programming," *IEEE/CAA J. Automa. Sinica,* vol. 4, no. 2, pp. 168–176, Apr. 2017.

[3] D. Liu, Y. Xu, Q. Wei, and X. Liu, "Residential energy scheduling for variable weather solar energy based on adaptive dynamic programming," *IEEE/CAA J. Automa. Sinica,* vol. 5, no. 1, pp. 36–46, Jan. 2018.

[4] D. Liu, S. Xue, B. Zhao, B. Luo, and Q. Wei, "Adaptive dynamic programming for control: a survey and recent advances," *IEEE Trans. Syst., Man, Cybern., Syst.,* vol. 51, no. 1, pp. 142–160, Jan. 2021.

[5] D. Liu and Q. Wei, "Policy iteration adaptive dynamic programming algorithm for discrete-time nonlinear systems," *IEEE Trans. Neural Netw. Learn Syst.,* vol. 25, no. 3, pp. 621–634, Mar. 2014.

[6] P. J. Werbos, "Approximate dynamic programming for real-time control and neural modeling," in *Handbook of Intelligent Control: Neural, Fuzzy, and Adaptive Approaches,* D. A. White and D. A. Sofge, Eds. New York, NY, USA: Van Nostrand Reinhold, 1992, ch. 13.

[7] B. Zhao, D. Liu, and Y. Li, "Observer based adaptive dynamic programming for fault tolerant control of a class of nonlinear systems" *Inf. Sci.,* vol. 384, pp. 21–33, 2017.

[8] B. Zhao, D. Liu, and C. Luo, "Reinforcement learning-based optimal stabilization for unknown nonlinear systems subject to inputs with uncertain constraints," *IEEE Trans. Neural Netw. Learn Syst.,* vol. 31, no. 10, pp. 4330–4340, Oct. 2020.

[9] D. Liu, D. Wang, and H. Li, "Decentralized stabilization for a class of continuous-time nonlinear interconnected systems using online learning optimal control approach," *IEEE Trans. Neural Netw. Learn Syst.,* vol. 25, no. 2, pp. 418–428, Feb. 2014.

[10] Y. Zhang, B. Zhao, D. Liu, and S. Zhang, "Event-triggered control of discrete-time zero-sum games via deterministic policy gradient adaptive dynamic programming," *IEEE Trans. Syst., Man, Cybern., Syst.,* doi: 10.1109/TSMC.2021.3105663.

[11] H. Zhang, H. Jiang, Y. Luo, and G. Xiao, "Data-driven optimal consensus control for discrete-time multi-agent systems with unknown dynamics using reinforcement learning method," *IEEE Trans. Ind. Electron.,* vol. 64, no. 5, pp. 4091–4100, May 2017.

[12] X. Yang, H. Zhang, and Z. Wang, "Data-based optimal consensus control for multiagent systems with policy gradient reinforcement learning," *IEEE Trans. Neural Netw. Learn Syst.,* doi: 10.1109/TNNLS.2021.3054685.

[13] J. Sun and T. Long, "Event-triggered distributed zero-sum differential game for nonlinear multi-agent systems using adaptive dynamic programming," *ISA Trans.* vol. 110, pp. 39–52, Apr. 2021.

[14] H. Wang and M. Li, "Model-free reinforcement learning for fully cooperative consensus problem of nonlinear multiagent systems," *IEEE Trans. Neural Netw. Learn Syst.,* doi: 10.1109/TNNLS.2020.3042508.

[15] S. Khankalantary, I. Izadi, and F. Sheikholeslam, "Robust ADP-based solution of a class of nonlinear multi-agent systems with input saturation and collision avoidance constraints" *ISA Trans.* vol. 107, pp. 52–62, Dec. 2020.

[16] Y. Jiang and Z. Jiang, "Robust adaptive dynamic programming and feedback stabilization of nonlinear systems," *IEEE Trans. Neural Netw. Learn Syst.,* vol. 25, no. 5, pp. 882–893, May 2014.

[17] D. Liu, X. Yang, D. Wang, and Q. Wei, "Reinforcement-learning-based robust controller design for continuous-time uncertain nonlinear systems subject to input constraints," *IEEE Trans. Cybern.,* vol. 45, no. 7, pp. 1372–1385, Jul. 2015.

[18] 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.

[19] D. Wang and D. Liu, "Learning and guaranteed cost control with event-based adaptive critic implementation," *IEEE Trans. Neural Netw. Learn Syst.,* vol. 29, no. 12, pp. 6004–6014, Dec. 2018.

[20] C. Mu, Y. Zhang, Z. Gao, and C. Sun, "ADP-based robust tracking control for a class of nonlinear systems with unmatched uncertainties," *IEEE Trans. Syst., Man, Cybern., Syst.,* vol. 50, no. 11, pp. 4056–4067, Nov. 2020.

[21] B. Zhao and D. Liu, "Event-triggered decentralized tracking control of modular reconfigurable robots through adaptive dynamic programming," *IEEE Trans. Ind. Electron.,* vol. 67, no. 4, pp. 3054–3064, Apr. 2020.

[22] B. Zhao, D. Liu, and C. Alippi, "Sliding-mode surface-based approximate optimal control for uncertain nonlinear systems with asymptotically stable critic structure," *IEEE Trans. Cybern.,* doi: 10.1109/TCYB.2019.2962011.

[23] M. Lin, B. Zhao, and D. Liu, "Policy gradient adaptive critic designs for model-free optimal tracking control with experience replay," *IEEE Trans. Syst., Man, Cybern., Syst.,* doi: 10.1109/TSMC.2021.3071968.

[24] B. Zhao, D. Wang, G. Shi, D. Liu, and Y. Li, "Decentralized control for large-scale nonlinear systems with unknown mismatched interconnections via policy iteration," *IEEE Trans. Syst., Man, Cybern., Syst.,* vol. 48, no. 10, pp. 1725–1735, Oct. 2018.

[25] B. Zhao, F. Luo, H. Lin, and D. Liu, "Particle swarm optimized neural networks based local tracking control scheme of unknown nonlinear interconnected systems," *Neural Netw.,* vol. 134, pp. 54–63, Feb. 2021.

[26] Y. Zhang, B. Zhao, and D. Liu, "Deterministic policy gradient adaptive dynamic programming for model-free optimal control," *Neurocomputing,* vol. 387, pp. 40–50, Jan. 2020.

[27] S. Zhang, B. Zhao, D. Liu, and Y. Zhang, "Observer-based event-triggered control for zero-sum games of input constrained multi-player nonlinear systems," *Neural Netw.,* vol. 144, pp. 101–112, Aug. 2021.