Automatica 148 (2023) 110768

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

Reinforcement learning and cooperative H_∞ output regulation of linear continuous-time multi-agent systems ☆

Yi Jiang [a,b], Weinan Gao [a], Jin Wu [c], Tianyou Chai [a,*], Frank L. Lewis [d]

[a] State Key Laboratory of Synthetical Automation for Process Industries and International Joint Research Laboratory of Integrated Automation, Northeastern University, Shenyang 110819, China
[b] Department of Electrical Engineering, City University of Hong Kong, Hong Kong Special Administrative Region of China
[c] Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong Special Administrative Region of China
[d] UTA Research Institute, University of Texas at Arlington, Fort Worth, TX 76118, USA

ARTICLE INFO

Article history:
Received 11 November 2020
Received in revised form 11 July 2022
Accepted 23 October 2022
Available online 12 December 2022

Keywords:
Cooperative H_∞ output regulation
Multi-agent system
Value iteration
Reinforcement learning

ABSTRACT

This paper proposes a novel control approach to solve the cooperative H_∞ output regulation problem for linear continuous-time multi-agent systems (MASs). Different from existing solutions to cooperative output regulation problems, a distributed feedforward-feedback controller is developed to achieve asymptotic tracking and reject both modeled and unmodeled disturbances. The feedforward control policy is computed via solving regulator equations, and the optimal feedback control policy is obtained through handling a zero-sum game. Instead of relying on the knowledge of system matrices in the state equations of the followers' dynamics and initial stabilizing feedback control gains, a value iteration (VI) algorithm is proposed to learn the optimal feedback control gain and feedforward control gain using online data. To the best of our knowledge, this paper is the first to show that the proposed VI algorithm can approximate the solution to continuous-time game algebraic Riccati equations with guaranteed convergence. Finally, the numerical analysis is provided to show the effectiveness of the proposed approach.

1. Introduction

The cooperative output regulation problem (CORP) for multi-agent systems (MASs) has attracted much attention in the last decade (Hu et al., 2018; Meng et al., 2015; Yang et al., 2014) due to its direct relevance to many real-world applications, such as spacecrafts, mobile robots, sensor networks. To solve the CORP for MASs, the designer needs to develop a servo-regulator for each follower to track the output of the leader (exosystem) under some communication typologies.

To achieve cooperative output regulation for MASs, one usually combines the techniques in the output regulation theory (Francis, 1977) and distributed observer design Su and Huang (2011), Yan and Huang (2016). In Su and Huang (2011) and Yan and Huang (2016), distributed observers were designed for linear continuous-time (CT) and discrete-time (DT) MASs, which were designed based on the knowledge of the system matrix of the leader. As an improvement, adaptive distributed observers were designed to estimate both the system matrix and state of the leader for linear CT and DT MASs to solve the CORP in Cai et al. (2017) and Huang (2017). However, most of these works assumed that the dynamics of each follower in MASs were precisely available.

Considering MASs with partially or completely unknown followers' dynamics, Ding (2017) developed adaptive controllers and estimated regulation errors of linear network systems under a digraph. In Zhang and Lewis (2012), a robust adaptive neural network controller was designed for each higher-order unknown nonlinear follower such that all follower nodes ultimately synchronized to the leader node with bounded residual errors. To ensure that the response of the closed-loop system is in an optimal sense with unknown follower dynamics, the reinforcement learning (RL) and adaptive dynamic programming (ADP) techniques have been introduced to learn an optimal regulator for unknown systems (Chen et al., 2022; Gao et al., 2022; Jiang, Fan et al., 2022; Jiang et al., 2019; Jiang, Fan, Chai, Lewis et al., 2018;

☆ This work was supported in part by NSFC 61991404, 61991403, 61991400, 2020 Science and Technology Major Project of Liaoning Province 2020JH1/10100008, 111 Project 2.0 (No. B08015). The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Raul Ordonez under the direction of Editor Miroslav Krstic.

* Corresponding author.
E-mail addresses: yjian22@cityu.edu.hk (Y. Jiang), gaown@mail.neu.edu.cn (W. Gao), jin_wu_uestc@hotmail.com (J. Wu), tychai@mail.neu.edu.cn (T. Chai), lewis@uta.edu (F.L. Lewis).

https://doi.org/10.1016/j.automatica.2022.110768
0005-1098/

Jiang, Fan, Chai, Li et al., 2018; Jiang, Gao et al., 2022; Kiumarsi et al., 2014; Pang et al., 2022; Xue et al., 2021, 2020). Given the reference trajectory, the optimal tracking problem for MASs was studied in Chen et al. (2020), Modares et al. (2016), Yang et al. (2018) and Zhang et al. (2017). Gao et al. (2019), Gao et al. (2018) and Gao et al. (2021) proposed RL-based approaches to achieve cooperative adaptive optimal output regulation of linear MASs. Qin et al. (2018) investigated the optimal synchronization problem for linear MASs with input saturation by using RL. For nonlinear MASs in DT, an RL-based approach proposed in Jiang, Fan et al. (2020) enabled the closed-loop system to achieve cooperative adaptive optimal output regulation.

In practice, MASs may be affected not only by modeled disturbances generated via an autonomous external system, but also some external unmodeled disturbances. The unmodeled disturbances are usually not studied in traditional output regulation problems. Yuan (2017) analyzed H_∞ output regulation of linear heterogeneous MASs to take care of the unmodeled disturbances. The H_∞ optimal control is related to zero-sum differential games with two players, which are the minimizing and maximizing players, respectively (Başar & Bernhard, 2008). However, most of existing solutions to (cooperative) H_∞ output regulation problems are model-based. Vamvoudakis and Lewis (2011) originally proposed an online adaptive control algorithm based on policy iteration (PI) RL to solve the CT multi-player non-zero-sum game. They further combined cooperative control, RL, and game theory to present a multi-agent formulation for the online solution of team games (Vamvoudakis et al., 2012). Fan et al. (2018), Li et al. (2021) and Rizvi and Lin (2018, 2020) proposed online adaptive control algorithms based on PI and VI RL to solve both the CT and DT H_∞ control problems using output feedback data. Nevertheless, to the best of our knowledge, the cooperative H_∞ output regulation problem (CH_∞ORP) with unknown system matrices in the state equations of the followers' dynamics has not been solved.

In this paper, we propose a novel data-driven approach to the CH_∞ORP for CT linear time-invariant (LTI) MASs and a feedforward-feedback controller is developed to achieve asymptotic tracking and reject both modeled and unmodeled disturbances. The technical contributions of this paper are as follows. As the first contribution, we have originally developed distributed controllers where the feedback control gains are obtained by solving game algebraic Riccati equations (GAREs), and the feedforward gains are found by solving regulator equations. Second, we have proposed a novel VI approach to solve the CT GARE using stochastic approximation methods (Bian & Jiang, 2015, 2016; Kushner & Yin, 2003). We have rigorously ensured the convergence of the proposed VI approach. Last but not least, without relying on the knowledge of system matrices in the state equations of the followers' dynamics, we are the first to solve the CH_∞ORP through proposing VI-based RL algorithms, which does not need an initial stabilizing control policy to start learning. Considering the fact that some followers are not able to access the leader's information, we build an adaptive distributed observer for followers to observe the dynamics and state of the leader to enable the cooperative output regulation by using VI-based RL algorithms.

The remainder of this paper is organized as follows. Section 2 formulates a CH_∞ORP. Section 3 provides the model-based solution to the CH_∞ORP and a model-based VI algorithm to approximate the optimal feedback control and worst disturbance gains, whose convergence proof is also shown therein. In Section 4, an online VI algorithm are presented to learn the optimal regulator for each follower. In Section 5, a simulation study is provided to show the effectiveness of the proposed method. Section 6 contains the concluding remarks and future work.

Notation: Throughout this paper, $\mathbb{R}$, $\mathbb{R}_+$ and $\mathbb{N}$ denote the sets of real numbers, positive real numbers and natural numbers excluding 0, respectively. For $n \in \mathbb{N}$ and matrices $X, Y \in \mathbb{R}^{n\times n}$, $X > Y$ ($X \geq Y$) means $X - Y$ is positive definite (positive semi-definite); X^{-1} denotes the inverse of nonsingular matrix X; $\sigma(X)$ is its complex spectrum. For brevity, denote $\mathbb{R}^n := \mathbb{R}^{n\times 1}$. Moreover, $\|\cdot\|$ denotes Euclidean norm for vectors or the Frobenius norm for matrices; $\otimes$ denotes Kronecker product; $\mathbb{P}^n$ denotes the space of all n-by-n real symmetric matrices; $\mathbb{P}^n_+ = \{P \in \mathbb{P}^n : P \geq 0\}$. For any $0 < T < \infty$, $\mathcal{F}([0, T], \mathbb{P}^n)$ denotes the space of functions from $[0, T]$ to $\mathbb{P}^n$. For $m, n \in \mathbb{N}$ and $X \in \mathbb{R}^{m\times n}$, X^{T} denotes the transpose of X; $0_{a\times b}$ denotes a null matrix of dimension $a \times b$ and I_a denotes an identity matrix of dimension $a\times a$; $\mathrm{vec}(X) = [x_1^{\mathrm{T}}, x_2^{\mathrm{T}}, \ldots, x_n^{\mathrm{T}}]^{\mathrm{T}}$ with $x_i \in \mathbb{R}^m$ being the columns of matrix X. For a symmetric matrix $X \in \mathbb{R}^{n\times n}$, $\mathrm{vecs}(X) = [x_{11}, 2x_{12}, \ldots, 2x_{1n}, x_{22}, 2x_{23}, \ldots, 2x_{(n-1)n}, x_{nn}]^{\mathrm{T}} \in \mathbb{R}^{(1/2)n(n+1)}$. For a column vector $\upsilon \in \mathbb{R}^n$, $\mathrm{vecv}(\upsilon) = [\upsilon_1^2, \upsilon_1\upsilon_2, \ldots, \upsilon_1\upsilon_n, \upsilon_2^2, \upsilon_2\upsilon_3, \ldots, \upsilon_{n-1}\upsilon_n, \upsilon_n^2]^{\mathrm{T}} \in \mathbb{R}^{(1/2)n(n+1)}$.

2. Problem formulation

This section presents the formulation of the CH_∞ORP of linear CT MASs. There are one leader and N followers in the concerned MAS. The leader generates the reference trajectory to be tracked and the external modeled disturbance to be rejected. Its dynamics is given as the following exosystem,

$$\dot{\upsilon}(t) = E\upsilon(t), \tag{1}$$

where $\upsilon \in \mathbb{R}^{n_\upsilon}$ and $E \in \mathbb{R}^{n_\upsilon\times n_\upsilon}$ is a constant matrix. The N followers are desired to track the reference trajectory and their state and output equations are given as the following disturbed CT LTI systems,

$$\dot{x}_i(t) = A_i x_i(t) + B_i u_i(t) + D_i w_i(t) + T_i \upsilon(t), \tag{2}$$

$$y_i(t) = C_i x_i(t) + S_i u_i(t), \; i = 1, \ldots, N, \tag{3}$$

where $x_i \in \mathbb{R}^{n_{xi}}$, $u_i \in \mathbb{R}^{n_{ui}}$, $w_i \in \mathbb{R}^{n_{wi}}$ and $y_i \in \mathbb{R}^{n_{yi}}$ are the state, input, unmodeled disturbance, and output of each follower, respectively. $A_i \in \mathbb{R}^{n_{xi}\times n_{xi}}$, $B_i \in \mathbb{R}^{n_{xi}\times n_{ui}}$, $D_i \in \mathbb{R}^{n_{xi}\times n_{wi}}$, $T_i \in \mathbb{R}^{n_{xi}\times n_\upsilon}$, $C_i \in \mathbb{R}^{n_{yi}\times n_{xi}}$ and $S_i \in \mathbb{R}^{n_{yi}\times n_{ui}}$ are constant matrices. Then, the tracking error $e_i \in \mathbb{R}^{n_{yi}}$ between each follower and the leader can be shown as follows,

$$e_i(t) := y_i(t) - y_{di}(t) = C_i x_i(t) + S_i u_i(t) + F_i \upsilon(t), \tag{4}$$

where $y_{di} = -F_i\upsilon \in \mathbb{R}^{n_{yi}}$ is the reference signal and $F_i \in \mathbb{R}^{n_{yi}\times n_\upsilon}$ is a constant matrix.

The connectivity of the leader and N followers is called as the communication topology of the MAS, which can be described using a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{\upsilon_0, \upsilon_1, \ldots, \upsilon_N\}$ is a set with υ_0 being the leader and υ_i, $i = 1, \ldots, N$ being N followers, and $\mathcal{E}$ is a set of edges with (υ_i, υ_j) as elements, which is termed an edge from υ_i to υ_j and means the agent i can use the state of the agent j. The adjacency matrix of the graph is $\mathcal{A} = [a_{ij}] \in \mathbb{R}^{N\times N}$ with $a_{ij} = 1$ if $(\upsilon_j, \upsilon_i) \in \mathcal{E}$, $a_{ij} = 0$ if $(\upsilon_j, \upsilon_i) \notin \mathcal{E}$ and $a_{ii} = 0$ for all $i = 1, \ldots, N$. Then, the Laplacian matrix $\mathcal{L}$ of graph $\mathcal{G}$ is given as follows,

$$\mathcal{L} = \begin{bmatrix} \sum_{j=1}^{N} a_{0j} & -[a_{01}, \ldots, a_{0N}] \\ -[a_{10}, \ldots, a_{N0}]^{\mathrm{T}} & H \end{bmatrix},$$

where $H = [h_{ij}] = \mathrm{diag}\{d_i(\upsilon_i)\} - \mathcal{A} \in \mathbb{R}^{N\times N}$ with $d_i(\upsilon_i) = \sum_{j=0}^{N} a_{ij}$ and $h_{ij} = -a_{ij}$, for any $i \neq j$, and $a_{i0} = 1$ if $(\upsilon_0, \upsilon_i) \in \mathcal{E}$ and $a_{i0} = 0$ if $(\upsilon_0, \upsilon_i) \notin \mathcal{E}$ and $a_{0i} = 0$ for all $i = 1, \ldots, N$.

For this MAS, the following assumptions are made.

Assumption 1. The pair (A_i, B_i), $i = 1, \ldots, N$ is stablizable.

Assumption 2. $\mathrm{rank}\left(\begin{bmatrix} A_i - \lambda I & B_i \\ C_i & S_i \end{bmatrix}\right) = n_{xi} + n_{yi}, \forall \lambda \in \sigma(E)$.

Assumption 3. All eigenvalues of E are semi-simple with zero real part.

Assumption 4. The directed graph $\mathcal{G}$ has a directed spanning tree with the leader being the root node.

Assumption 5. The unmodeled disturbance satisfies $w_i(t) \in \mathcal{L}_2[0,\infty)$ for each follower.

Remark 1. Assumptions 1–3 are standard in the linear output regulation problem and ensure its solvability (Huang, 2004). Under Assumption 3, $v(t)$ is excited when $v(0) \neq 0$. Assumption 4 is standard in the cooperative control of MASs subject to fixed networks. Under this assumption, H is a nonsingular matrix (Su & Huang, 2011) and all the nonzero eigenvalues of H, if any, have positive real parts. Assumption 5 is standard in the H_∞ control problem.

The CH$_\infty$ORP is solved if one can design regulators for all followers such that: (1) the cooperative output regulation is achieved; (2) the closed-loop systems is robust to the unmodeled disturbance $w_i(t)$ in the sense of $\mathcal{L}_2$. The control input for each follower is expressed as follows,

$$u_i(t) = -K_i x_i(t) + L_i \eta_i(t), \tag{5}$$

where $K_i \in \mathbb{R}^{n_{ui}\times n_{xi}}$ is the feedback control gain and $L_i \in \mathbb{R}^{n_{ui}\times n_v}$ is the feedforward control gain. $\eta_i(t) \in \mathbb{R}^{n_v}$ is the observation of the leader's state $v(t)$ for the ith follower. Inspired by Cai et al. (2017), an adaptive distributed observer for each follower to observe $v(t)$ and E is designed as follows,

$$\dot{E}_i(t) = \mu_1 \sum_{j=0}^{N} a_{ij}(E_j(t) - E_i(t)), \tag{6}$$

$$\dot{\eta}_i(t) = E_i(t)\eta_i(t) + \mu_2 \sum_{j=0}^{N} a_{ij}(\eta_j(t) - \eta_i(t)), \tag{7}$$

where $E_i(t) \in \mathbb{R}^{n_v\times n_v}$ is the observed values of E, $\eta_0(t) = v(t)$ and $E_0(t) = E$. The following lemma shows how to design constants μ_1 and μ_2 so that the observed errors $\tilde{E}_i(t) := E_i(t) - E$ and $\tilde{\eta}_i(t) := \eta_i(t) - v(t)$ are asymptotically stable.

Lemma 1 (*Cai et al., 2017, Lemma 2*)**.** *Given the leader system in (1) and adaptive distributed observers in (6)–(7), under Assumption 4, for any initial $\eta_i(t_0)$ and $E_i(t_0)$, $i = 1,\dots,N$, one has that: (1) for $\mu_1 > 0$, $\lim_{t\to\infty} \tilde{E}_i(t) = 0$ exponentially; (2) for sufficiently large $\mu_1, \mu_2 > 0$, $\lim_{t\to\infty} \tilde{\eta}_i(t) = 0$ exponentially.*

To accommodate the requirement (1) of the CH$_\infty$ORP, the (cooperative) output regulation theory (Francis, 1977; Huang, 2004) is usually adopted. The following theorem shows how to solve the CORP by solving regulator equations. Before proceeding, we provide the following lemma.

Lemma 2 (*Cai et al., 2017, Lemma 1*)**.** *Consider the system $\dot{x}(t) = Fx(t) + F_1(t)$ where $x \in \mathbb{R}^n$, $F \in \mathbb{R}^{n\times n}$ is Hurwitz, $F_1(t) \in \mathbb{R}^n$ is bounded for all $t > t_0$. Then, one has that if $\lim_{t\to\infty} F_1(t) = 0$, for any $x(t_0)$, $\lim_{t\to\infty} x(t) = 0$.*

Theorem 1. *Consider the MAS in (2)–(4). Under Assumptions 1–5, K_i are designed to assure that $A_i - B_iK_i$ are Hurwitz. The CORP is solved by controllers in (5) if L_i is computed as*

$$L_i = U_i + K_i X_i, \tag{8}$$

where matrices $X_i \in \mathbb{R}^{n_{xi}\times n_{vi}}$ and $U_i \in \mathbb{R}^{n_{ui}\times n_{vi}}$ are solutions to the following regulator equations,

$$X_i E = A_i X_i + B_i U_i + T_i, \tag{9}$$

$$0 = C_i X_i + S_i U_i + F_i. \tag{10}$$

Proof. Assumption 2 implies that regulator Eqs. (9)–(10) are solvable (Huang, 2004). By defining $\bar{x}_i(t) := x_i(t) - X_i v(t)$, from (1)–(4), (8)–(10) and the definition of $\eta_i(t)$, one has

$$\dot{\bar{x}}_i(t) = (A_i - B_iK_i)\bar{x}_i(t) + D_i w_i(t) + B_i L_i \tilde{\eta}_i(t). \tag{11}$$

Under Assumption 5 and Lemmas 1–2, since $A_i - B_iK_i$ are Hurwitz, one has that $\lim_{t\to\infty} \bar{x}_i(t) = 0$, which implies $\lim_{t\to\infty} e_i(t) = \lim_{t\to\infty}[(C_i - S_iK_i)\bar{x}_i(t) + S_iL_i\tilde{\eta}_i(t)] = 0$ and completes the proof of Theorem 1. □

The CORP focuses on how to find a proper feedback control gain K_i and a proper feedforward control gain L_i. Under the solutions X_i and U_i to the regulator Eqs. (9)–(10) and the accessible $v(t)$, a new disturbed CT LTI system can be formulated as

$$\dot{\bar{x}}_i(t) = A_i\bar{x}_i(t) + B_i\bar{u}_i(t) + D_i w_i(t), \tag{12}$$

$$e_i(t) = C_i\bar{x}_i(t) + S_i\bar{u}_i(t), \tag{13}$$

with $\bar{u}_i(t) := u_i(t) - U_i v(t)$. To solve the CH$_\infty$ORP in the presence of $w_i(t)$, which requires that the system (11) is robust to the unmodeled disturbance $w_i(t)$ in the sense of $\mathcal{L}_2$, inspired by Krener (1992), we need to solve two separated optimization problems. The first one, Problem 1, is a static constrained optimization problem and solved to obtain a unique pair (X_i, U_i). While the second one, Problem 2, is a dynamic constrained minimax problem and solved to obtain the optimal feedback control gain K_i^*. Problems 1–2 are formulated as follows:

Problem 1. Solve the following static constrained optimization problem to find a unique pair (X_i, U_i) for $i = 1,\dots,N$ satisfying regulator Eqs. (9)–(10),

$$\begin{aligned} \min_{\bar{X}_i} \quad & \bar{X}_i^{\mathrm{T}} M_i \bar{X}_i \\ \text{s.t.} \quad & (9)\text{–}(10), \end{aligned} \tag{14}$$

where $\bar{X}_i = \mathrm{vec}([X_i^{\mathrm{T}}, U_i^{\mathrm{T}}]^{\mathrm{T}})$, $M_i > 0$ is set by users. ■

Problem 2. Solve the following dynamic constrained minimax problem for $i = 1,\dots,N$,

$$\begin{aligned} \min_{\bar{u}_i}\max_{w_i} \quad & V_i(\bar{x}_i(0); \bar{u}_i, w_i) = \int_0^\infty r_i(\tau)d\tau \\ \text{s.t.} \quad & (12) \end{aligned} \tag{15}$$

to obtain the optimal feedback control policy

$$\bar{u}_i^*(t) = -K_i^*\bar{x}_i(t) \tag{16}$$

and the worst disturbance policy

$$w_i^*(t) = W_i^*\bar{x}_i(t), \tag{17}$$

where $r_i(\tau) := \bar{x}_i^{\mathrm{T}}(\tau)Q_i\bar{x}_i(\tau) + \bar{u}_i^{\mathrm{T}}(\tau)R_i\bar{u}_i(\tau) - \gamma_i^2 w_i^{\mathrm{T}}(\tau)w_i(\tau)$ with $Q_i > 0$ and $R_i > 0$. γ_i is a prescribed constant disturbance attenuation level satisfying $\gamma_i > \gamma_i^* > 0$. ■

The control problem that this paper aims to solve is presented as follows. Under the Assumptions 1–5, when the system matrices A_i, B_i, D_i and T_i in the state Eqs. (2) are unknown, the system matrices C_i, S_i and F_i in the error Eqs. (4) are known, solve Problems 1–2 for each follower by using online information $x_i(t)$, $u_i(t)$, $w_i(t)$ and $\eta_i(t)$.

Remark 2. Problem 2 is a standard H_∞ control problem and solvable if and only if $\gamma_i > \gamma_i^* := \inf\{\gamma_i > 0 \mid \min_{\bar{u}_i}\max_{w_i} V_i(0; \bar{u}_i, w_i) \le 0\}$. Therefore, since the system matrices A_i, B_i and D_i are unknown, one can select γ_i as large as possible so that the data-driven approach proposed in the following section can be adopted. By solving this problem, one can obtain a stabilizing feedback control gain K_i^* for each follower, and the corresponding worst disturbance policy (17) can be represented as $w_i^*(t) = W_i^* x_i(t) - W_i^* X_i v(t)$.

Remark 3. Problem 2 considers the H_∞ performance of (12) instead of (11) with respect to $w_i(t)$ due to that $\tilde{E}_i(0) = 0$ and $\tilde{\eta}_i(0) = 0$, which is a zero initial value requirement and standard in the H_∞ control problem. Under these conditions and $\bar{x}_i(0) = 0$, for any disturbance $w_i(t)$ satisfying Assumption 5, the closed-loop system of (12) has $\mathcal{L}_2$-gain less than or equal to γ_i (Başar & Bernhard, 2008; van der Schaft, 1992). Then, the control performance $\sum_{i=1}^N \int_0^\infty (\bar{x}_i^T(\tau) Q_i \bar{x}_i(\tau) + \bar{u}_i^T(\tau) R_i \bar{u}_i(\tau)) d\tau \le \sum_{i=1}^N \int_0^\infty \gamma_i^2 w_i^T(\tau) w_i(\tau) d\tau$ is thus satisfied. In the rest of this paper, it will be shown in Theorem 4 that the solution to this problem will result in a closed-loop system which is robust to $w_i(t)$ and $\tilde{\eta}_i(t)$ in the sense of $\mathcal{L}_2$ when $\tilde{\eta}_i(t)$ is also considered as an external disturbance.

3. Solution to the CH_∞ORP with known dynamics

In this section, we present solutions to Problems 1–2 by assuming that all followers' dynamics are known and the leader's dynamics and state are available for all followers. For Problem 1, Lagrange multiplier is introduced to convert this problem to a static unconstrained optimization problem and hence the solution to this problem can be provided. A novel model-based VI algorithm is provided to obtain the optimal solution to Problem 2. Results shown in this section will be helpful to develop off-policy VI algorithms to compute the optimal feedback control gain K_i^* and the optimal feedforward control gain L_i^* online for each follower in Section 4.

3.1. Model-based solutions of regulator equations

To deal with Problem 1, inspired by Fan et al. (2020), Gao and Jiang (2016) and Jiang, Kiumarsi et al. (2020), two Sylvester maps $\Omega_i : \mathbb{R}^{n_{xi} \times n_v} \to \mathbb{R}^{n_{xi} \times n_v}$ and $\bar{\Omega}_i : \mathbb{R}^{n_{xi} \times n_v} \times \mathbb{R}^{n_{ui} \times n_v} \to \mathbb{R}^{n_{xi} \times n_v}$ for $i = 1, \ldots, N$ are introduced, they are

$$\Omega_i(X_i) := X_i E - A_i X_i, \tag{18}$$

$$\bar{\Omega}_i(X_i, U_i) := X_i E - A_i X_i - B_i U_i. \tag{19}$$

Under these two Sylvester maps, general solutions of (10) can be established. For $i = 1, \ldots, N$, select two sequences $X_{il} \in \mathbb{R}^{n_{xi} \times n_v}$ and $U_{il} \in \mathbb{R}^{n_{ui} \times n_v}$ with $l = 0, 1, \ldots, m_i$, where $(m_i - 1)$ equals to the dimension of the null space of $I_{n_v} \otimes [C_i, S_i]$, $X_{i0} = 0_{n_{xi} \times n_v}$, $U_{i0} = 0_{n_{ui} \times n_v}$, X_{i1} and U_{i1} are selected to satisfy $C_i X_{i1} + S_i U_{i1} = -F_i$, and all the vectors $\mathrm{vec}([X_{il}^T, U_{il}^T]^T)$ form a basis for $\ker(I_{n_v} \otimes [C_i, S_i])$ with $l = 2, 3, \ldots, m_i$, that is, $C_i X_{il} + S_i U_{il} = 0$. Clearly, (X_i, U_i) must be the summation of (X_{i1}, U_{i1}) and a linear combination of (X_{il}, U_{il}), $l = 2, 3, \ldots, m_i$. General solutions of (10) can be established as

$$(X_i, U_i) = (X_{i1}, U_{i1}) + \sum_{l=2}^{m_i} \alpha_{il} (X_{il}, U_{il}),$$

where $\alpha_{il} \in \mathbb{R}$ are some unknown coefficients. Besides, Eq. (9) can be established as

$$\bar{\Omega}_i(X_i, U_i) = \bar{\Omega}_i(X_{i1}, U_{i1}) + \sum_{l=2}^{m_i} \alpha_{il} \bar{\Omega}_i(X_{il}, U_{il}) = T_i.$$

Thus, regulator Eqs. (9)–(10) can be rewritten as the following equation,

$$\Lambda_i \chi_i = \xi_i, \quad i = 1, \ldots, N, \tag{20}$$

where

$$\Lambda_i = \begin{bmatrix} \mathrm{vec}(\bar{\Omega}_i(X_{i2}, U_{i2})) & \cdots & \mathrm{vec}(\bar{\Omega}_i(X_{im_i}, U_{im_i})) & 0 \\ \mathrm{vec}\left(\left[X_{i2}^T, U_{i2}^T\right]^T\right) & \cdots & \mathrm{vec}\left(\left[X_{im_i}^T, U_{im_i}^T\right]^T\right) & -I \end{bmatrix},$$

$$\chi_i = \begin{bmatrix} \alpha_{i2} & \cdots & \alpha_{im_i} & \bar{X}_i^T \end{bmatrix}^T,$$

$$\xi_i = \begin{bmatrix} \mathrm{vec}(-\bar{\Omega}_i(X_{i1}, U_{i1}) + T_i) \\ -\mathrm{vec}\left(\left[X_{i1}^T, U_{i1}^T\right]^T\right) \end{bmatrix} =: \begin{bmatrix} \xi_{i1} \\ \xi_{i2} \end{bmatrix}.$$

Then, by using row operation, Eq. (20) can be rewritten as

$$\begin{bmatrix} \bar{\Lambda}_{i11} & \bar{\Lambda}_{i12} \\ \bar{\Lambda}_{i21} & \bar{\Lambda}_{i22} \end{bmatrix} \chi_i = \begin{bmatrix} \bar{\xi}_{i1} \\ \bar{\xi}_{i2} \end{bmatrix} \tag{21}$$

with $\bar{\Lambda}_{i21} \in \mathbb{R}^{(m_i-1)\times(m_i-1)}$ being nonsingular matrices for all i. Above equations can be rewritten as

$$\Pi_i \bar{X}_i = \Psi_i, \tag{22}$$

where $\Pi_i = -\bar{\Lambda}_{i11} \bar{\Lambda}_{i21}^{-1} \bar{\Lambda}_{i22} + \bar{\Lambda}_{i12}$, $\Psi_i = -\bar{\Lambda}_{i11} \bar{\Lambda}_{i21}^{-1} \bar{\xi}_{i2} + \bar{\xi}_{i1}$.

By using the Lagrange multiplier method, Problem 1 can be rewritten as static unconstrained optimization problems as

$$\min_{\bar{X}_i} \quad J_i = \bar{X}_i^T M_i \bar{X}_i + \lambda_i^T (\Pi_i \bar{X}_i - \Psi_i), \tag{23}$$

where λ_i is the Lagrange multiplier. Setting the partial derivative of the objective function J_i in (23) with respect to $\bar{X}_i$ along with the constraint conditions (22) equal to 0, one has that optimal solutions of Problem 1 satisfy the following equations

$$\begin{bmatrix} 2M_i & \Pi_i^T \\ \Pi_i & 0 \end{bmatrix} \begin{bmatrix} \bar{X}_i \\ \lambda_i \end{bmatrix} = \begin{bmatrix} 0 \\ \Psi_i \end{bmatrix}. \tag{24}$$

3.2. Model-based VI algorithm for CH_∞ORP

Problem 2 is a standard H_∞ control problem. By Chen (2000), the minimax problem in (15) can be solved via the optimal feedback control gain

$$K_i^* = R_i^{-1} B_i^T P_i^* \tag{25}$$

and the worst disturbance gain

$$W_i^* = \gamma_i^{-2} D_i^T P_i^* \tag{26}$$

for $i = 1, \ldots, N$, where P_i^* is the unique positive definite solution to the following GARE ensuring that the matrix $A_i - B_i K_i^* + D_i W_i^*$ is Hurwitz,

$$A_i^T P_i + P_i A_i + Q_i - P_i (B_i R_i^{-1} B_i^T - \gamma_i^{-2} D_i D_i^T) P_i = 0. \tag{27}$$

Clearly, the GARE is a nonlinear function on P_i. An alternative way to solve this GARE is iteration based approach, such as PI algorithm. In the PI based approach (Jiang, Zhang et al., 2022; Odekunle et al., 2020), initial stabilizing feedback control and disturbance gains are required simultaneously to achieve convergence, which is a very strict condition due to A_i, B_i and D_i being unknown for regulators design. In the VI based approach, this requirement can be relaxed. Algorithm 1 shows how to use model-based VI approach to solve the GARE in (27) and hence the CH_∞ORP can be solved. Theorem 2 analyzes the convergence of the proposed model-based VI algorithm. Before introducing Algorithm 1, define collection of bounded sets $\{B_q^i\}_{q=0}^\infty$ with non-empty interiors and a sequence $\{\epsilon_j\}_{j=0}^\infty$ satisfying

$$B_q^i \subseteq B_{q+1}^i, q \in \mathbb{N}, \lim_{q\to\infty} B_q^i = \mathbb{P}_+^{n_{xi}},$$

$$\epsilon_j > 0, \sum_{j=0}^\infty \epsilon_j = \infty, \sum_{j=0}^\infty \epsilon_j^2 < \infty.$$

Algorithm 1 Model-based VI Algorithm for CH_∞ORP

Input: A_i, B_i, D_i, T_i, C_i, S_i, E
Output: P_i^*, K_i^*, W_i^*, L_i^*

$K_{i0} \leftarrow 0$, $W_{i0} \leftarrow 0$, $P_{i0} \leftarrow 0$, $j \leftarrow 0$, $q \leftarrow 0$, $i \leftarrow 1$, $m \leftarrow 0$
Select a small threshold $\varepsilon > 0$ as stopping criterion
repeat
  $i \leftarrow 1$, $m \leftarrow m + 1$
  repeat
    $j \leftarrow 0$, $q \leftarrow 0$
    loop
      $\tilde{P}_{i(j+1)} = P_{ij} + (0.5)^m \epsilon_j (A_i^{\mathrm{T}} P_{ij} + P_{ij} A_i + Q_i - K_{ij}^{\mathrm{T}} R_i K_{ij} + \gamma_i^2 W_{ij}^{\mathrm{T}} W_{ij})$, $K_{i(j+1)} = R_i^{-1} B_i^{\mathrm{T}} \tilde{P}_{i(j+1)}$, $W_{i(j+1)} = \gamma_i^{-2} D_i^{\mathrm{T}} \tilde{P}_{i(j+1)}$
      if $\tilde{P}_{i(j+1)} \notin B_q^i$ **then**
        $P_{i(j+1)} \leftarrow P_{i0}$, $q \leftarrow q + 1$
      if $\|\tilde{P}_{i(j+1)} - P_{ij}\|/\epsilon_j < \varepsilon$ **then**
        return $P_{ij^*} = \tilde{P}_{i(j+1)}$, $K_{ij^*} = K_{ij}$ and $W_{ij^*} = W_{ij}$ as the approximations of P_i^*, K_i^* and W_i^*
      else
        $P_{i(j+1)} \leftarrow \tilde{P}_{i(j+1)}$, $j \leftarrow j + 1$
    end loop
    $i \leftarrow i + 1$
  until $i = N + 1$
until $Q_i + K_{ij^*}^{\mathrm{T}} R_i K_{ij^*} - \gamma_i^2 W_{ij^*}^{\mathrm{T}} W_{ij^*} > 0$
Solve (X_i, U_i) using (24)
Calculate L_i^* using (8)

Theorem 2. *Consider $\{P_{ij}\}_{j=0}^{\infty}$ and $\{K_{ij}\}_{j=0}^{\infty}$ in Algorithm 1. Under Assumptions 1–3, for $i = 1, \ldots, N$, if $\gamma_i > \gamma_i^*$ and the threshold ε is set as 0, $\lim_{j\to\infty} P_{ij} = P_i^*$, $\lim_{j\to\infty} K_{ij} = K_i^*$ and (X_i, U_i) solves the regulator Eqs. (9)–(10).*

Proof. See the Appendix. □

Remark 4. The convergence analysis of VI Algorithm 1 for CT systems is technically challenging. To begin with, one has to show that the solution of a game differential Riccati equation converges to that of a GARE as the time goes to $-\infty$. Inspired from Kučera (1973), we show this fact in Lemma 5 based on a sufficient condition that $\gamma_i > \gamma_i^*$. Then, we show that P_{ij} will stay in a predefined compact set for any large enough iteration j in Lemma 6. By Lemmas 5–6, we finally show the convergence of the proposed VI Algorithm 1. Besides, an alternative way to prove the convergence is also provided.

Remark 5. The standard approach to find the unique positive definite solution P_i^* to the GARE is to reduce to the problem of solving a generalized eigenvalue problem. However, the Riccati problem may be ill-conditioned, and the resulting solution may not be as accurate as desired. The detail discusses can be found in Arnold and Laub (1984). The iteration-based approach in Algorithm 1 is a widely applicable approach to avoid this problem and easy to be implemented. Besides, compared with the model-based PI approach (Jiang, Zhang et al., 2022; Odekunle et al., 2020), Algorithm 1 has two advantages. First, the initial stabilizing feedback control and disturbance gains are not required to start the iteration, which means $A_i - B_i K_{i0} + D_i W_{i0}$ should not be Hurwitz for $i = 1, \ldots, N$. Second, there is no need to solve the Lyapunov equation at each iteration but only simple iterative computation is needed.

Remark 6. Since P_i^* is the unique and minimal positive definite solution of the GARE in (27) enabling $A_i - B_i K_i^* + D_i W_i^*$ being Hurwitz (Başar & Bernhard, 2008, Section 9, Theorem 9.7), if $A_i - B_i K_{ij^*} + D_i W_{ij^*}$ is Hurwitz, P_{ij^*} can be proven to converge to P_i^*. This condition is satisfied when $Q_i + K_{ij^*}^{\mathrm{T}} R_i K_{ij^*} - \gamma_i^2 W_{ij^*}^{\mathrm{T}} W_{ij^*} > 0$. If this term is negative definite, smaller $\{\epsilon_j\}_{j=0}^{\infty}$ is required to make sure that P_{ij} will converge to the minimal positive definite solution P_i^*.

The model-based solutions provided in this section can be calculated based on two assumptions. The first one is that the leader's state $v(t)$ and the matrix E in (1) are accessible to all followers, which contradicts to the communication topology in this work. Another one is that the system matrices A_i, B_i, D_i and T_i are known so that Algorithm 1 can be applied. In the next section, we will remove these two requirements by using online data and show how the Algorithm 1 can help us to implement the online data-driven approach.

4. Data-driven solution for the CH_∞ORP

In this section, we propose an online iteration algorithm to learn the solution of the CH_∞ORP by using the online data. This section includes three parts: (1) the optimal feedback control gain K_i^* and worst disturbance gain W_i^* online computation; (2) the optimal feedforward control gain L_i^* online computation; (3) the algorithm realization and performance analysis.

4.1. VI based solution to the GARE

By defining $\tilde{x}_{il}(t) := x_i(t) - X_{il}\eta_i(t)$ for $i = 1, \ldots, N$ and $l = 0, 1, \ldots, m_i$, one has that

$$\begin{aligned}\dot{\tilde{x}}_{il}(t) =& A_i x_i(t) + B_i u_i(t) + D_i w_i(t) + T_i v(t) - X_{il} E \eta_i(t) \\ &+ X_{il} \tilde{E}_i(t) \eta_i(t) - \mu_2 X_{il} \sum_{j=0}^{N} a_{ij}(\eta_j(t) - \eta_i(t)). \end{aligned} \tag{28}$$

Taking the derivative of $\tilde{x}_{il}^{\mathrm{T}}(t) P_{ij} \tilde{x}_{il}(t)$ yields

$$\begin{aligned}& d(\tilde{x}_{il}^{\mathrm{T}}(t) P_{ij} \tilde{x}_{il}(t))/dt \\ =& \tilde{x}_{il}^{\mathrm{T}}(t) H_{ij} \tilde{x}_{il}(t) + 2u_i^{\mathrm{T}}(t) R_i K_{ij} \tilde{x}_{il}(t) + 2\gamma_i^2 \tilde{x}_{il}^{\mathrm{T}}(t) W_{ij} w_i(t) \\ &- 2\tilde{x}_{il}^{\mathrm{T}}(t) G_{ijl} \eta_i(t) + 2\tilde{x}_{il}^{\mathrm{T}}(t) P_{ij} \rho_{il}(t), \end{aligned} \tag{29}$$

where $H_{ij} = A_i^{\mathrm{T}} P_{ij} + P_{ij} A_i$, $G_{ijl} = P_{ij}[T_i - \Omega_i(X_{il})]$, $K_{ij} = R_i^{-1} B_i^{\mathrm{T}} P_{ij}$, $W_{ij} = \gamma_i^{-2} D_i^{\mathrm{T}} P_{ij}$ and $\rho_{il}(t) = X_{il} \tilde{E}_i(t) \eta_i(t) - \mu_2 X_{il} \sum_{j=0}^{N} a_{ij}(\eta_j(t) - \eta_i(t)) - T_i \tilde{\eta}_i(t)$. Integrating both sides of above equation along the trajectories of (29), one has

$$\begin{aligned}& \tilde{x}_{il}^{\mathrm{T}}(t+\delta t) P_{ij} \tilde{x}_{il}(t+\delta t) - \tilde{x}_{il}^{\mathrm{T}}(t) P_{ij} \tilde{x}_{il}(t) \\ =& \int_t^{t+\delta t} \tilde{x}_{il}^{\mathrm{T}}(\tau) H_{ij} \tilde{x}_{il}(\tau) d\tau + \int_t^{t+\delta t} 2u_i^{\mathrm{T}}(\tau) R_i K_{ij} \tilde{x}_{il}(\tau) d\tau \\ &+ \int_t^{t+\delta t} 2\gamma_i^2 \tilde{x}_{il}^{\mathrm{T}}(\tau) W_{ij} w_i(\tau) d\tau + \int_t^{t+\delta t} 2\tilde{x}_{il}^{\mathrm{T}}(\tau) P_{ij} \rho_{il}(\tau) d\tau \\ &+ \int_t^{t+\delta t} 2\tilde{x}_{il}^{\mathrm{T}}(\tau) G_{ijl} \eta_i(\tau) d\tau, \end{aligned} \tag{30}$$

where $\delta > 0$. Next, for a positive integer s, define following matrices:

$$\begin{aligned}\xi_{\tilde{x}_{il}\tilde{x}_{il}} =& \left[\mathrm{vecv}(\tilde{x}_{il}(t_1)) - \mathrm{vecv}(\tilde{x}_{il}(t_0)), \ldots, \right. \\ & \left. \mathrm{vecv}(\tilde{x}_{il}(t_s)) - \mathrm{vecv}(\tilde{x}_{il}(t_{s-1}))\right]^{\mathrm{T}}, \\ \Gamma_{\tilde{x}_{il}\tilde{x}_{il}} =& \left[\int_{t_0}^{t_1} \mathrm{vecv}(\tilde{x}_{il}(\tau)) d\tau, \ldots, \int_{t_{s-1}}^{t_s} \mathrm{vecv}(\tilde{x}_{il}(\tau)) d\tau\right]^{\mathrm{T}}, \\ \Gamma_{\tilde{x}_{il}u_i} =& 2\left[\int_{t_0}^{t_1} \tilde{x}_{il}(\tau) \otimes R_i u_i(\tau) d\tau, \ldots, \int_{t_{s-1}}^{t_s} \tilde{x}_{il}(\tau) \otimes R_i u_i(\tau) d\tau\right]^{\mathrm{T}}, \end{aligned}$$

$$\Gamma_{w_i\tilde{x}_{il}} = 2\gamma_i^2\left[\int_{t_0}^{t_1} w_i(\tau)\otimes\tilde{x}_{il}(\tau)d\tau,\ldots,\int_{t_{s-1}}^{t_s} w_i(\tau)\otimes\tilde{x}_{il}(\tau)d\tau\right]^{\mathrm{T}},$$

$$\Gamma_{\eta_i\tilde{x}_{il}} = 2\left[\int_{t_0}^{t_1} \eta_i(\tau)\otimes\tilde{x}_{il}(\tau)d\tau,\ldots,\int_{t_{s-1}}^{t_s} \eta_i(\tau)\otimes\tilde{x}_{il}(\tau)d\tau\right]^{\mathrm{T}},$$

$$\Gamma_{\rho_{il}\tilde{x}_{il}} = 2\left[\int_{t_0}^{t_1} \rho_{il}(\tau)\otimes\tilde{x}_{il}(\tau)d\tau,\ldots,\int_{t_{s-1}}^{t_s} \rho_{il}(\tau)\otimes\tilde{x}_{il}(\tau)d\tau\right]^{\mathrm{T}},$$

where $t_0 < t_1 < \cdots < t_s$ are positive. Using above matrices, one can rewrite (30) as

$$\begin{aligned}&\Gamma_{il}\left[\mathrm{vecs}(H_{ij})^{\mathrm{T}}, \mathrm{vec}(K_{ij})^{\mathrm{T}}, \mathrm{vec}(W_{ij})^{\mathrm{T}}, \mathrm{vec}(G_{ijl})^{\mathrm{T}}\right]^{\mathrm{T}}\\ =&\xi_{\tilde{x}_{il}\tilde{x}_{il}}\mathrm{vecs}(P_{ij}) - \Gamma_{\rho_{il}\tilde{x}_{il}}\mathrm{vec}(P_{ij}),\end{aligned} \tag{31}$$

where $\Gamma_{il} = \left[\Gamma_{\tilde{x}_{il}\tilde{x}_{il}}, \Gamma_{\tilde{x}_{il}u_i}, \Gamma_{w_i\tilde{x}_{il}}, \Gamma_{\eta_i\tilde{x}_{il}}\right]$. In (31), one needs small positive constants ε_1 and ε_2 to estimate the convergence of $\rho_{il}(t)$. Based on Lemma 1, if $\bar{\mu}_{i1}(\bar{t}) := \|\sum_{j=0}^{N} a_{ij}(E_j(\bar{t}) - E_i(\bar{t}))\| < \varepsilon_1$ and $\bar{\mu}_{i2}(\bar{t}) := \|\sum_{j=0}^{N} a_{ij}(\eta_j(\bar{t}) - \eta_i(\bar{t}))\| < \varepsilon_2$ hold, there exists a small constant ε_3 such that $\|\rho_{il}(t)\| < \varepsilon_3, \forall t > \bar{t}$ and ε_3 will be sufficiently small as ε_1 and ε_2 are chosen to be sufficiently small. Eq. (31) can be solved by using least squares. The uniqueness of solution to (31) is guaranteed under the following excitation assumption, which is made in some related works, see Lee et al. (2014, (29)) and Vamvoudakis and Lewis (2010, Remark 2).

Assumption 6. There exist $\underline{\beta}_{il} > 0$ and $\bar{\beta}_{il} > 0$ and proper control input $u_i(t)$, $t \in [\bar{t}, t_s)$ such that $\underline{\beta}_{il}I_{n_{ri}} \leq \Gamma_{il}^{\mathrm{T}}\Gamma_{il} \leq \bar{\beta}_{il}I_{n_{ri}}$, where $n_{ri} = \frac{n_{xi}(n_{xi}+1)}{2} + n_{xi}(n_{ui} + n_{wi} + n_v)$.

From the statement in Lemma 1, $\lim_{t\to\infty}\rho_{il}(t) = 0$ asymptotically and hence $\lim_{t\to\infty}\Gamma_{\rho_{il}\tilde{x}_i}(t) = 0$ asymptotically. Under Assumption 6, one has the solutions of (31), they are

$$\begin{aligned}&\left[\mathrm{vecs}(H_{ij})^{\mathrm{T}}, \mathrm{vec}(K_{ij})^{\mathrm{T}}, \mathrm{vec}(W_{ij})^{\mathrm{T}}, \mathrm{vec}(G_{ijl})^{\mathrm{T}}\right]^{\mathrm{T}}\\ =&\Theta_{il}\xi_{\tilde{x}_{il}\tilde{x}_{il}}\mathrm{vecs}(P_{ij}), \text{ as } t\to\infty,\end{aligned} \tag{32}$$

where $\Theta_{il} = \left(\Gamma_{il}^{\mathrm{T}}\Gamma_{il}\right)^{-1}\Gamma_{il}^{\mathrm{T}}$. Note that by using (32), the optimal feedback control and worst disturbance gains computation in Algorithm 1 can be realized without using A_i, B_i, D_i and T_i.

4.2. Optimal feedforward control gain computation

From Eq. (32), one can obtain K_{ij} and G_{ijl} for $i = 1,\ldots,N$ and $l = 1,\ldots,m_i$. Then, one can calculate B_i, T_i and $\Omega_i(X_{il})$ by $B_i = P_{ij}^{-1}K_{ij}^{\mathrm{T}}R_i$, $T_i = P_{ij}^{-1}G_{ij0}$ and $\Omega_i(X_{il}) = -P_{ij}^{-1}G_{ijl} + P_{ij}^{-1}G_{ij0}$, respectively. As a result, Λ_i and ξ_i in (20) become

$$\Lambda_i = \begin{bmatrix}\mathrm{vec}(-P_{ij}^{-1}G_{ij2} + P_{ij}^{-1}G_{ij0} - P_{ij}^{-1}K_{ij}^{\mathrm{T}}R_iU_{i2}) & \cdots & \mathrm{vec}(-P_{ij}^{-1}G_{ijm_i} + P_{ij}^{-1}G_{ij0} - P_{ij}^{-1}K_{ij}^{\mathrm{T}}R_iU_{im_i}) & 0\\ \mathrm{vec}\left(\left[X_{i2}^{\mathrm{T}}, U_{i2}^{\mathrm{T}}\right]^{\mathrm{T}}\right) & \cdots & \mathrm{vec}\left(\left[X_{im_i}^{\mathrm{T}}, U_{im_i}^{\mathrm{T}}\right]^{\mathrm{T}}\right) & -I\end{bmatrix},$$

$$\xi_i = \begin{bmatrix}\mathrm{vec}(P_{ij}^{-1}G_{ij1} + P_{ij}^{-1}K_{ij}^{\mathrm{T}}R_iU_{i1})\\ -\mathrm{vec}\left(\left[X_{i1}^{\mathrm{T}}, U_{i1}^{\mathrm{T}}\right]^{\mathrm{T}}\right)\end{bmatrix}, \tag{33}$$

and hence (X_i, U_i) can be solved by using (24). Based on (X_i, U_i), the optimal feedforward control gain can be calculated by using (8).

Algorithm 2 Online VI Algorithm for the CH$_\infty$ORP

Input: Θ_{il}, $\xi_{\tilde{x}_{il}\tilde{x}_{il}}$, X_{il} and U_{il} with $l = 0, 1, \ldots, m_i$
Output: P_i^*, K_i^*, W_i^*, L_i^*

$K_{i0} \leftarrow 0$, $W_{i0} \leftarrow 0$, $P_{i0} \leftarrow 0$, $j \leftarrow 0$, $q \leftarrow 0$, $i \leftarrow 1$, $m \leftarrow 0$
Select a small threshold $\varepsilon > 0$ as stopping criterion
repeat
  $i \leftarrow 1$, $m \leftarrow m + 1$
  repeat
    $j \leftarrow 0$,$q \leftarrow 0$
    loop
      Solve for H_{ij}, K_{ij} and W_{ij} using (32), $\tilde{P}_{i(j+1)} = P_{ij} + (0.5)^m\epsilon_j(H_{ij} + Q_i - K_{ij}^{\mathrm{T}}R_iK_{ij} + \gamma_i^2W_{ij}^{\mathrm{T}}W_{ij})$
      if $\tilde{P}_{i(j+1)} \notin B_q$ **then**
        $P_{i(j+1)} \leftarrow P_{i0}$, $q \leftarrow q + 1$
      if $\|\tilde{P}_{i(j+1)} - P_{ij}\|/\epsilon_j < \varepsilon_i$ **then**
        return $P_{ij^*} = \tilde{P}_{i(j+1)}$, $K_{ij^*} = K_{ij}$ and $W_{ij^*} = W_{ij}$ as the approximations of P_i^*, K_i^* and W_i^*, where $j^* = j$
      else
        $P_{i(j+1)} \leftarrow \tilde{P}_{i(j+1)}$, $j \leftarrow j + 1$
    end loop
    $i \leftarrow i + 1$
  until $i = N + 1$
until $Q_i + K_{ij^*}^{\mathrm{T}}R_iK_{ij^*} - \gamma_i^2W_{ij^*}^{\mathrm{T}}W_{ij^*} > 0$
$i \leftarrow 1$, $l \leftarrow 0$
repeat
  repeat
    Solve G_{ijl} by using (32) with $P_{ij} = P_{ij^*}$, $l \leftarrow l + 1$
  until $l = m_i + 1$
  $i \leftarrow i + 1$, $l \leftarrow 0$
until $i = N + 1$
Solve Λ_i and ξ_i for $i = 1,\ldots,N$ via (33)
Obtain Π_i and Ψ_i for $i = 1,\ldots,N$ via (22)
Solve (X_i, U_i) for $i = 1,\ldots,N$ via (24)
Calculate L_i^* using (8)

4.3. Algorithm realization and performance analysis

Finally, an online VI based algorithm is provided for solving the CH$_\infty$ORP as Algorithm 2. In Algorithm 2, $u_i(t)$ is employed as the input on $[\bar{t}, t_s]$ to ensure Assumption 6. Then, we compute $\xi_{\tilde{x}_{il}\tilde{x}_{il}}$, $\Gamma_{\tilde{x}_{il}\tilde{x}_{il}}$, $\Gamma_{\tilde{x}_{il}u_i}$, $\Gamma_{w_i\tilde{x}_{il}}$, $\Gamma_{\eta_i\tilde{x}_{il}}$ to run Algorithm 2. By using Algorithm 2, one can obtain the earned feedback control gain K_{ij^*} and learned feedforward control gain L_i^*. The CH$_\infty$ORP is hence solved by using the following controller of each follower on $(t_s, \infty]$,

$$u_i(t) = -K_{ij^*}x_i(t) + L_i^*\eta_i(t), \forall t > t_s. \tag{34}$$

Besides, the following two theorems are provided to analyze the convergence of this algorithm and performance of the resulting closed-loop MAS.

Theorem 3. *Consider $\{P_{ij}\}_{j=0}^{\infty}$ and $\{K_{ij}\}_{j=0}^{\infty}$ obtained via Algorithm 2. Under Assumptions 1–6, for $i = 1,\ldots,N$, if $\gamma_i > \gamma_i^*$, ε, ε_1 and ε_2 are sufficiently small, μ_1 and μ_2 are sufficiently large, one has that $\lim_{\varepsilon_1,\varepsilon_2\to0,j\to\infty}P_{ij} = P_i^*$, $\lim_{\varepsilon_1,\varepsilon_2\to0,j\to\infty}K_{ij} = K_i^*$ and (X_i, U_i) solves regulator equations.*

Proof. Under Assumption 6, Eq. (32) has a unique solution for $i = 1,\ldots,N$. Denoting $\hat{H}_{ij}$, $\hat{K}_{ij}$ and $\hat{W}_{ij}$ as the solutions to Eq. (32). Then, without losing generality, we assume that $\Theta_{il}\Gamma_{\rho_{il}\tilde{x}_{il}}\mathrm{vec}(P_{ij}) = [\mathrm{vecs}(\Delta_1(P_{ij}))^{\mathrm{T}}, \mathrm{vec}(\Delta_2(P_{ij}))^{\mathrm{T}}, \mathrm{vec}(\Delta_3(P_{ij}))^{\mathrm{T}}, \mathrm{vec}(\Delta_4(P_{ij}))^{\mathrm{T}}]^{\mathrm{T}}$. Then, based on the statement in Remark 6, one obtains $\lim_{\varepsilon_1,\varepsilon_2\to0}[\mathrm{vecs}(\Delta_1(P_{ij}))^{\mathrm{T}}, \mathrm{vec}(\Delta_2(P_{ij}))^{\mathrm{T}}, \mathrm{vec}(\Delta_3(P_{ij}))^{\mathrm{T}},$

$\mathrm{vec}(\Delta_4(P_{ij}))^{\mathrm{T}}] = 0$. Therefore, $\hat{H}_{ij} = H_{ij} + \Delta_1(P_{ij})$, $\hat{K}_{ij} = K_{ij} + \Delta_2(P_{ij})$ and $\hat{W}_{ij} = W_{ij} + \Delta_3(P_{ij})$. This implies that $\tilde{P}_{i(j+1)}$ in Algorithm 2 can be viewed as

$$\hat{\tilde{P}}_{i(j+1)} = \hat{P}_{ij} + \epsilon_j(\hat{H}_{ij} + Q_i - \hat{K}_{ij}^{\mathrm{T}} R \hat{K}_{ij} + \gamma_i^2 \hat{W}_{ij}^{\mathrm{T}} \hat{W}_{ij}).$$

Under the statement in Theorem 2, it then follows from $\lim_{\varepsilon_1,\varepsilon_2\to 0}[\mathrm{vecs}(\Delta_1(P_{ij}))^{\mathrm{T}}, \mathrm{vec}(\Delta_2(P_{ij}))^{\mathrm{T}}, \mathrm{vec}(\Delta_3(P_{ij}))^{\mathrm{T}}, \mathrm{vec}(\Delta_4(P_{ij}))^{\mathrm{T}}] = 0$ that $\lim_{\varepsilon_1,\varepsilon_2\to 0, j\to\infty} \hat{P}_{ij} = P_i^*$, $\lim_{\varepsilon_1,\varepsilon_2\to 0, j\to\infty} \hat{K}_{ij} = K_i^*$. Moreover, as $\varepsilon_1, \varepsilon_2 \to 0$, Λ_i and ξ_i in (33) are the same as them in (20). Then, (X_i, U_i) solved via Algorithm 2 are equivalent to the solutions solved via Algorithm 1, which completes the proof of Theorem 3. □

Theorem 4. *Consider the MAS in (1)–(3). Under Assumptions 1–6, with the learned feedback control gain K_{ij^*} and learned feedforward control gain L_i^* obtained by Algorithm 2, the tracking errors of all followers satisfy $\lim_{t\to\infty} e_i(t) = 0$. Moreover,*

$$\begin{aligned} &0.5\int_0^\infty \left(\bar{x}_i^{\mathrm{T}}(t) Q_i \bar{x}_i(t) + \bar{u}_i^{\mathrm{T}}(t) R_i \bar{u}_i(t)\right) dt \\ &\leq \int_0^\infty \left(\gamma_i^2 \|w_i(t)\|^2\right) dt + V_i^{\odot}(\bar{x}_i(0), \tilde{\eta}_i(0)), \end{aligned} \tag{35}$$

where $V_i^{\odot}(\bar{x}_i(t), \tilde{\eta}_i(t)) = \bar{x}_i^{\mathrm{T}}(t) P_i^{\odot} \bar{x}_i(t) + \int_t^\infty \left(2\|R_i^{\frac{1}{2}} L_i^{\odot} \tilde{\eta}_i(\tau)\|^2\right) d\tau$, $P_i^{\odot}$ and $L_i^{\odot}$ are the solutions solved via Algorithm 2. That is, the closed-loop system is robust to the unmodeled disturbance $w_i(t)$ and $\tilde{\eta}_i(t)$ in the sense of $\mathcal{L}_2$.

Proof. There exist small enough ε_i such that

$$A_i^{\mathrm{T}} P_i^{\odot} + P_i^{\odot} A_i + 0.5 Q_i < P_i^{\odot}(B_i R_i^{-1} B_i^{\mathrm{T}} - \gamma_i^{-2} D_i D_i^{\mathrm{T}}) P_i^{\odot}.$$

The inequalities above can be rewritten by

$$\begin{aligned} &(A_i - B_i K_{ij^*})^{\mathrm{T}} P_i^{\odot} + P_i^{\odot}(A_i - B_i K_{ij^*}) + 0.5 Q_i + K_{ij^*}^{\mathrm{T}} R_i K_{ij^*} \\ &+ \gamma_i^{-2} P_i^{\odot} D_i D_i^{\mathrm{T}} P_i^{\odot} < 0. \end{aligned}$$

Taking the derivative for $V_i^{\odot}(\bar{x}_i(t), \eta_i(t))$ around the solutions of the error systems (11) yields

$$\begin{aligned} \dot{V}_i^{\odot} \leq &- 0.5\bar{x}_i^{\mathrm{T}}(t) Q_i \bar{x}_i(t) - \bar{u}_i^{\mathrm{T}}(t) R_i \bar{u}_i(t) - \|\gamma_i D_i^{\mathrm{T}} P_i^{\odot} \bar{x}_i(t)\|^2 \\ &+ 2w_i^{\mathrm{T}}(t) D_i^{\mathrm{T}} P_i^{\odot} \bar{x}_i(t) + 2\bar{u}_i^{\mathrm{T}}(t) R_i L_i^{\odot} \tilde{\eta}_i(t) - 2\|R_i^{\frac{1}{2}} L_i^{\odot} \tilde{\eta}_i(t)\|^2 \\ \leq &- 0.5\bar{x}_i^{\mathrm{T}}(t) Q_i \bar{x}_i(t) - 0.5\bar{u}_i^{\mathrm{T}}(t) R_i \bar{u}_i(t) + \gamma_i^2 \|w_i(t)\|^2 \\ &- \|\gamma_i D_i^{\mathrm{T}} P_i^{\odot} \bar{x}_i(t) - \gamma_i w_i(t)\|^2. \end{aligned} \tag{36}$$

Integrating (36) from $t = 0$ to $t = \infty$, one has

$$\begin{aligned} &\int_0^\infty \left(0.5\bar{x}_i^{\mathrm{T}}(t) Q_i \bar{x}_i(t) + 0.5\bar{u}_i^{\mathrm{T}}(t) R_i \bar{u}_i(t)\right) dt \\ &\leq \int_0^\infty \left(\gamma_i^2 \|w_i(t)\|^2\right) dt + V_i^{\odot}(\bar{x}_i(0), \tilde{\eta}_i(0)) - \lim_{t\to\infty} V_i^{\odot}(\bar{x}_i(t), \tilde{\eta}_i(t)) \\ &\leq \int_0^\infty \left(\gamma_i^2 \|w_i(t)\|^2\right) dt + V_i^{\odot}(\bar{x}_i(0), \tilde{\eta}_i(0)). \end{aligned} \tag{37}$$

Through Barbalat's lemma, one has $\lim_{t\to\infty} \bar{u}_i(t) = 0$ and $\lim_{t\to\infty} \bar{x}_i(t) = 0$. It is immediate to have $\lim_{t\to\infty} e_i(t) = 0$, which completes the proof of Theorem 4. □

Remark 7. In Eq. (32), $\Theta_{il}\xi_{\tilde{x}_{il}\tilde{x}_{il}}$ just need to be computed only once for each i and l since they just include online data. Besides, different from the online computation for K_i^* and W_i^*, in the online computation for L_i^*, Eq. (32) needs to be computed only once to obtain G_{ijl} for each i and l, which is shown as follows: define that $\Theta_{il}\xi_{\tilde{x}_{il}\tilde{x}_{il}} = [Z_{i1l}^{\mathrm{T}}, Z_{i2l}^{\mathrm{T}}]^{\mathrm{T}}$, where Z_{i1l} is the first $\frac{n_{xi}(n_{xi}+1)}{2} + n_{xi}(n_{ui}+n_{wi})$ rows of the matrix $\Theta_{il}\xi_{\tilde{x}_{il}\tilde{x}_{il}}$ and Z_{i2l} is the rest rows of the matrix $\Theta_{il}\xi_{\tilde{x}_{il}\tilde{x}_{il}}$. Then, one can only compute (H_{ij}, K_{ij}, W_{ij}) via $\left[\mathrm{vecs}(H_{ij})^{\mathrm{T}}, \mathrm{vec}(K_{ij})^{\mathrm{T}}, \mathrm{vec}(W_{ij})^{\mathrm{T}}\right]^{\mathrm{T}} = Z_{i1l}\mathrm{vecs}(P_{ij})$.

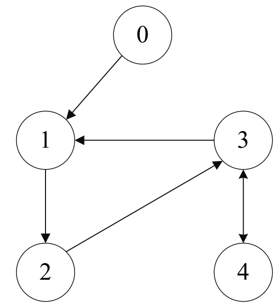


Fig. 1. Communication graph $\mathcal{G}$.

Remark 8. The online computation in Algorithm 2 depends on the convergence of the distributed observer. It follows from Lemma 1 that for arbitrary small positive constants ε_1 and ε_2, $\bar{t}$ can be small enough by choosing sufficiently large μ_1 and μ_2. However, there will always be a non-zero term $\rho_{il}(t)$ no matter how small it is. Therefore, in practical application, P_{ij} and K_{ij} obtained in Algorithm 2 converge to the sufficient small neighbors of P_i^* and K_i^* by choosing sufficiently small ε_1 and ε_2.

5. Numerical analysis

In this section, we provide a numerical analysis to illustrate the effectiveness of the proposed approach for the CH$_\infty$ORP. Consider the following MAS, the leader system is given as

$$\dot{v}(t) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} v(t), \quad y_{di}(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} v(t).$$

The 4 followers' systems are

$$\begin{aligned} \dot{x}_i(t) &= \begin{bmatrix} 0 & 1 \\ -1 & -3 \end{bmatrix} x_i(t) + \begin{bmatrix} 0 \\ i \end{bmatrix} u_i(t) + \begin{bmatrix} -i \\ i \end{bmatrix} w_i(t) + iv(t), \\ y_i(t) &= \begin{bmatrix} 1 & 0 \end{bmatrix} x_i(t) + u_i(t), \end{aligned}$$

with the disturbance inputs being $w_i(t) = \exp(-0.1t)(0.2\cos(12it) + \cos(6it)\sin(7it))$. Clearly, $w_i(t) \in \mathcal{L}_2[0, \infty)$. The communication graph $\mathcal{G}$ of this MAS is shown in Fig. 1 and same as it in Jiang, Fan et al. (2020).

In this section, we select the parameters in Problems 1 and 2 as $M_i = I$, $Q_i = I$, $R_i = 1$, $\gamma_i = 5$, $i = 1, \ldots, 4$. In Algorithm 2, the parameters are selected as follows: $B_q^i = \{X > 0 \mid \|X\| < 30(q+1)\}$, $\bar{t} = 5$ s, $t_s = 15$ s, for $i = 1, \ldots, 4$, $\varepsilon = 0.01$, $\epsilon_j = 2/j$. The control input on $[0, t_s]$ is chosen as $u_i(t) = 0.2\sin(15t/i) + \sin(5t)\sin(8t)$ to ensure Assumption 6. Initial values are selected as $x_i(0) = [1, 1]^{\mathrm{T}}$ and $v(0) = [1, 1]^{\mathrm{T}}$.

Based on these parameters, the simulation experiment result is shown. Fig. 2 shows trajectories of outputs of the leader and followers. It can be observed that all followers track the leader via learned feedback control and feedforward control gains by Algorithm 2. Fig. 3 shows trajectories of $\|P_{ij} - P_i^*\|$, which illustrates the conclusion in Theorem 3. It can be observed that the proposed data-driven approach can learn the optimal feedback control and feedforward control gains for the CH$_\infty$ORP by using online data of the MAS without using system matrices in the state equations of the followers' dynamics.

6. Conclusion

In this paper, we have studied the CH$_\infty$ORP in the presence of unknown system matrices in the state equations of the followers'

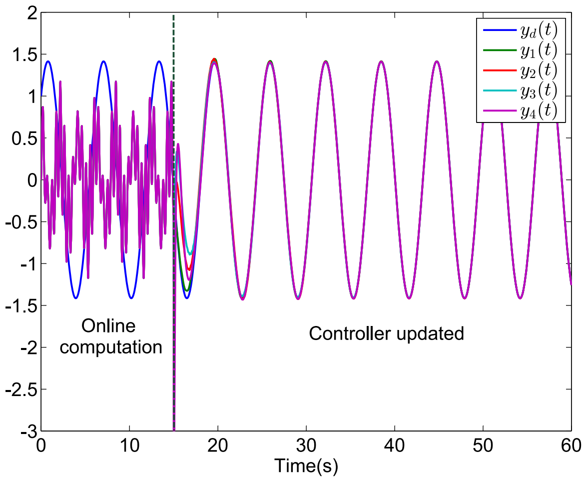


Fig. 2. Trajectories of outputs of the leader and followers.

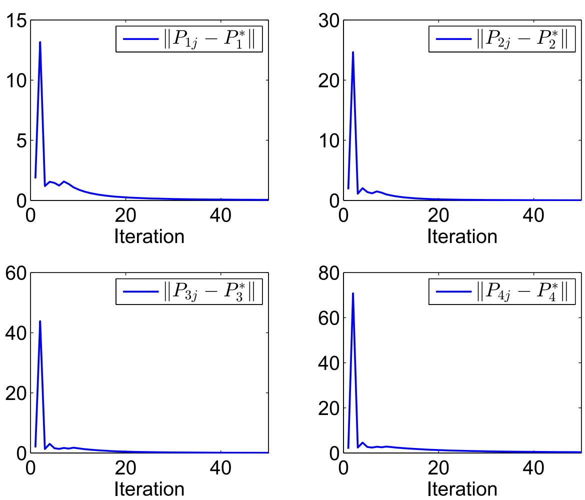


Fig. 3. Trajectories of $\|P_{ij} - P_i^*\|$.

dynamics with external disturbances. Under the requirements on cooperative output regulation and control performance, without relying on the knowledge of system matrices in the state equations of the followers' dynamics and initial stabilizing feedback control gains, an adaptive distributed observer and a VI based CH$_\infty$ORP algorithm is designed, which is capable of learning the optimal feedback control gain, worst disturbance gain and optimal feedforward control gain for each follower using online data and the observed value of the leader. The future work will focus on solving the CH$_\infty$ORP via output-feedback.

Appendix

Firstly, we should provide preliminary lemmas.

Lemma 3 (*Başar & Bernhard, 2008; Başar & Olsder, 1998, Section 4*)**.** *Consider the following ordinary differential equation (ODE),*

$$-\dot{P}_i(t) = A_i^{\mathrm{T}} P_i(t) + P_i(t) A_i + Q_i - P_i(t)(B_i R_i^{-1} B_i^{\mathrm{T}} - \gamma_i^{-2} D_i D_i^{\mathrm{T}}) P_i(t),\ P_i(t_f) \geq 0. \tag{A.1}$$

Under Assumption 1, for $i = 1, \ldots, N$, if $\gamma_i > \gamma_i^{CL} := \inf\{\gamma_i > 0 |\}$(A.1) does not have a conjugate point on $[t_0, t_f]$, $P_i(t)$ in (A.1) does not have a conjugate point on $[t_0, t_f]$.

Lemma 4 (*Başar & Bernhard, 2008, Section 9*)**.** *For any interval $[t_0, t_f]$, $\gamma_i^* \geq \gamma_i^{CL}$.*

Lemma 5. *Consider the ODEs (A.1). Under Assumption 1, for $i = 1, \ldots, N$, if $\gamma_i > \gamma_i^*$, $P_i(t)$ is bounded on $(-\infty, 0]$, $\lim_{t\to-\infty} P_i(t) = P_i^*$ for $P_i(0) = 0$ and P_i^* is the local asymptotically stable equilibrium point.*

Proof. Denote that

$$\bar{x}_i^{\mathrm{T}}(t_0) f_i(t_0, P, t_1) \bar{x}_i(t_0) = \min_{\bar{u}_i} \max_{w_i} \left[\bar{x}_i^{\mathrm{T}}(t_1) P \bar{x}_i(t_1) + \int_{t_0}^{t_1} r_i(\tau) d\tau \right].$$

From Başar and Bernhard (2008, Section 4, Theorem 4.1), $f_i(t_0, P_i(t_f), t_f) : \mathbb{R} \times \mathbb{R}^{n_{xi} \times n_{xi}} \times \mathbb{R} \to \mathbb{R}^{n_{xi} \times n_{xi}}$ are the solutions to the ODEs (A.1) at t_0. Since $\gamma_i > \gamma_i^*$, based on the conclusions in Lemmas 3–4, it follows from Assumption 1 that for any $t_0 \leq t_1 \leq t_2$, $\bar{x}_i^{\mathrm{T}}(t_0) f_i(t_0, 0, t_1) \bar{x}_i(t_0) \leq \bar{x}_i^{\mathrm{T}}(t_0) f_i(t_0, 0, t_2) \bar{x}_i(t_0)$, which implies that $f_i(t_0, 0, t_1) \leq f_i(t_0, 0, t_2)$ and for any $t_1 \leq t_2 \leq t_f$, $f_i(t_2, 0, t_f) \leq f_i(t_1, 0, t_f)$. Then, one can conclude that $f_i(t, 0, t_f)$ is monotonically nondecreasing with respect to t in the ordering of positive matrices for $i = 1, \ldots, N$. Moreover, one has $\bar{x}_i^{\mathrm{T}}(t_0) f_i(t_0, 0, t_f) \bar{x}_i(t_0) \leq \min_{\bar{u}_i} \max_{w_i} \int_{t_0}^{\infty} r_i(\tau) d\tau \leq \infty$, which implies that $f_i(t_0, 0, t_f)$ is bounded for $i = 1, \ldots, N$. Therefore, there exist finite matrices $P_{i\infty}$ such that $\lim_{t\to-\infty} P_i(t) = P_{i\infty}$. Since $P_{i\infty}$ is finite and independent of t for $i = 1, \ldots, N$ and solves the GARE in (27), one has that $P_{i\infty}$ is a positive definite solution of (27). Let us consider the case that $0 < P_i(0) \leq P_i^*$. From the definition of $\bar{x}_i^{\mathrm{T}}(t_0) f_i(t_0, P_i(0), t_1) \bar{x}_i(t_0)$, it will suffice to show that the term $x_i^{\mathrm{T}}(t_1) P_i(0) x_i(t_1)$ approaches zero as $t_1 \to \infty$. This term goes either to zero or to infinity. The latter case is not possible due to the boundedness. Then, one has that $\lim_{t\to-\infty} P_i(t)$ exists. Since $f_i(t_0, P_i(0), t_1) \geq f_i(t_0, 0, t_1)$, one can conclude that $f_i(t_0, 0, t_1)$ will converge to the minimal positive definite solution of (27). From Başar and Bernhard (2008, Section 9, Theorem 9.7), P_i^* is the minimal positive definite solution of (27). Therefore, one has that $\lim_{t\to-\infty} P_i(t) = P_i^*, \forall 0 \leq P_i(0) \leq P_i^*$, which completes the proof of the local attractivity of P_i^*. The rest of the proof will prove that P_i^* is the local asymptotically stable equilibrium point. Since $f_i(t, 0, t_0)$ is monotonically nondecreasing with respect to t, for each $\bar{\varepsilon}_i > 0$ and $P_i(t_0) \leq P_i^*$, one can always find a $P_i(t_0)$ satisfying $\|P_i(t_0) - P_i^*\| < \delta_i(\bar{\varepsilon}_i)$ such that $\|P_i(t) - P_i^*\| < \bar{\varepsilon}_i, \forall t \leq t_0$, which completes the proof of the local asymptotic stability of P_i^* and thus the proof of Lemma 5 is completed. □

Lemma 6. *Consider $\{P_{ij}\}_{j=0}^{\infty}$ and $\{K_{ij}\}_{j=0}^{\infty}$ in Algorithm 1. Under Assumption 1, for $i = 1, \ldots, N$, if $\gamma_i > \gamma_i^*$, one has that there exist $N_i \geq 0$ and a compact set $\mathcal{K}_i \subset \mathbb{P}_+^{n_{xi}}$, with nonempty interior, such that $\{P_{ij}\}_{j=0}^{\infty} \subset \mathcal{K}_i$.*

Proof. Let us reverse the timeline in (A.1) to obtain the following ODEs,

$$\dot{P}_i(t) = A_i^{\mathrm{T}} P_i(t) + P_i(t) A_i + Q_i - P_i(t)(B_i R_i^{-1} B_i^{\mathrm{T}} - \gamma_i^{-2} D_i D_i^{\mathrm{T}}) P_i(t),\ P_i(0) = 0. \tag{A.2}$$

Form the conclusion in Lemma 5, $\lim_{t\to\infty} P_i(t) = P_i^*$ asymptotically. By denoting $p_i = \mathrm{vecs}(P_i)$, we have the following ODEs,

$$\dot{p}_i = h(p_i), \tag{A.3}$$

where $h(p_i) = \mathrm{vecs}(A_i^{\mathrm{T}} P_i(t) + P_i(t) A_i + Q_i - P_i(t)(B_i R_i^{-1} B_i^{\mathrm{T}} - \gamma_i^{-2} D_i D_i^{\mathrm{T}}) P_i(t))$. Clearly, for $i = 1, \ldots, N$, $\lim_{t\to\infty} p_i(t) = p_i^* = \mathrm{vecs}(P_i^*)$ asymptotically with the initial $p_i(0)$ in the region of attraction of p_i^*, where $0 \leq P_i(0) \leq P_i^*$ and the region is denoted as R_{Ai}. Therefore, based on Khalil (1996, Theorem 4.17), there exists a smooth function $\bar{V}_i(p_i) : R_{Ai} \to \mathbb{R}_+$ for $i = 1, \ldots, N$, such that $\nabla\bar{V}_i(p_i)h(p_i) < 0$, $\bar{V}_i(p_i) > 0$, $\forall p_i \in R_{Ai}\backslash\{p_i^*\}$, $\nabla\bar{V}_i(p_i^*)h(p_i^*) = 0$ and $\bar{V}_i(p_i^*) = 0$. Obviously, for all $0 < C < \bar{V}_i(0)$, there exists a compact set $\{p_i|\bar{V}_i(p_i) < C, p_i \in R_{Ai}\}$ being a subset of R_{Ai}. When $P_{ij} \neq P_i^*$, one can always find two positive C_0 and C_1 so that $C_0 < \bar{V}_i(p_{ij}) < C_1$, where $p_{ij} = \text{vecs}(P_{ij})$. Choose

$$\delta_i \in (0, \inf_{\{p_i|C_0<\bar{V}_i(p_i)<C_1, p_i\in R_{Ai}\}} |\nabla\bar{V}_i(p_i)h(p_i)|).$$

Then, there exists $\bar{\epsilon}_{i1} > 0$, such that for all $0 < \epsilon < \bar{\epsilon}_{i1}$, $t \in [0, 1]$ and $p_i \in \Xi_i = \{p_i|C_0 < \bar{V}_i(p_i) < C_1, p_i \in R_{Ai}\}$, one has that $p_i + \epsilon th(p_i) \in R_{Ai}$ and

$$|\nabla\bar{V}_i(p_i)h(p_i) - \nabla\bar{V}_i(p_i + \epsilon th(p_i))h(p_i)| \le \inf_{\Xi_i}|\nabla\bar{V}_i(p_i)h(p_i)| - \delta_i.$$

The above equation implies that

$$\begin{aligned}&\bar{V}_i(p_i + \epsilon h(p_i)) - \bar{V}_i(p_i)\\ =&\epsilon\nabla\bar{V}_i(p_i)h(p_i) + \epsilon\int_0^1(\nabla\bar{V}_i(p_i + \epsilon th(p_i))h(p_i) - \nabla\bar{V}_i(p_i)h(p_i))dt\\ \le& -\epsilon\inf_{\Xi_i}|\nabla\bar{V}_i(p_i)h(p_i)| + \epsilon\left(\inf_{\Xi_i}|\nabla\bar{V}_i(p_i)h(p_i)| - \delta_i\right) = -\epsilon\delta_i.\end{aligned}$$

Therefore, for all $p_i \in \Xi_i$, $\bar{V}_i(p_i + \epsilon h(p_i)) < \bar{V}_i(p_i) < C_1$. If $\bar{V}_i(p_i) < C_0$, there exists $\bar{\epsilon}_{i2} > 0$, such that $\bar{V}_i(p_i + \epsilon h(p_j)) < C_1$, $\forall\epsilon < \bar{\epsilon}_{i2}$, $\forall\bar{V}_i(p_i) < C_0$. Now, by denoting $N_i := \inf\{j \ge \bar{N}_i|P_{ij} = P_{i0}\}$, where $\bar{N}_i \in \mathbb{N}$, one has that $\{P_{ij}\}_{j=0}^{\infty} \subset \mathcal{K}_i$, $\forall j \ge N_i$, where $\mathcal{K}_i = \mathbb{P}_+^{n_{xi}} \cap \{P_i|\bar{V}_i(\text{vecs}(P_i)) < C_1\}$. Since $\{P_i|\bar{V}_i(\text{vecs}(P_i)) < C_1\}$ is a compact set, $\mathcal{K}_i$ is a compact set. The proof of Lemma 6 is completed. □

Then, we are in position to present the proof of Theorem 2.

Proof of Theorem 2. The proof of the first part of this Theorem is similar to the proofs in Bian and Jiang (2016, Theorem 3.3) and Kushner and Yin (2003, Chapter 5). From the conclusion in Lemma 6, one can rewrite the iteration equations in Algorithm 1 as

$$P_{i(j+1)} = \tilde{P}_{i(j+1)} + Z_{ij}, \forall j > N_i, \tag{A.4}$$

where the projection terms Z_{ij} are defined as

$$Z_{ij} = \begin{cases} P_{i0} - \tilde{P}_{i(j+1)}, & \text{if } \tilde{P}_{i(j+1)} \notin \mathcal{K}_i,\\ 0, & \text{otherwise.}\end{cases}$$

Consider the following CT interpolations:

$$P_i^0(t) = \begin{cases} P_{i0}, & t \le 0,\\ P_{ij}, & t \in [t_j, t_{j+1}),\end{cases}$$

where $t_0 = 0$ and $t_j = \sum_{k=0}^{j-1}\epsilon_k$. Define the shifted process $P_i^j(t) = P_i^0(t_j + t)$, $\forall t \in (-\infty, \infty)$. Then, one has that

$$\begin{aligned}P_i^j(t) =&P_i^0(t_j + t) = P_{ip(t_j+t)}\\ =&P_{ij} + \sum_{k=j}^{p(t_j+t)-1}\epsilon_k(A_i^{\mathrm{T}}P_k + P_kA_i + Q_i\\ &- P_k(B_iR_i^{-1}B_i^{\mathrm{T}} - \gamma_i^{-2}D_iD_i^{\mathrm{T}})P_k) + \sum_{k=j}^{p(t_j+t)-1}Z_k\\ :=&P_{ij} + H_i^j(t) + e_i^j(t) + Z_i^j(t),\end{aligned} \tag{A.5}$$

where

$$\begin{aligned}H_i^j(t) =& \int_0^t(A_i^{\mathrm{T}}P_i^j(s) + P_i^j(s)A_i + Q_i\\ &- P_i^j(s)(B_iR_i^{-1}B_i^{\mathrm{T}} - \gamma_i^{-2}D_iD_i^{\mathrm{T}})P_i^j(s))ds,\\ e_i^j(t) =& \sum_{k=j}^{p(t_j+t)-1}\epsilon_k(A_i^{\mathrm{T}}P_k + P_kA_i + Q_i\\ &- P_k(B_iR_i^{-1}B_i^{\mathrm{T}} - \gamma_i^{-2}D_iD_i^{\mathrm{T}})P_k) - H_i^j(t),\\ Z_i^j(t) =& \sum_{k=j}^{p(t_j+t)-1}Z_k, \; p(t) = \begin{cases}k, & 0 \le t_k \le t \le t_{k+1}\\ 0, & t < 0.\end{cases}\end{aligned}$$

Based on the conclusion in Lemma 6, similar to the proof in Bian and Jiang (2016, Theorem 3.3), $\{H_i^j(\cdot)\}_{j=N_i}^{\infty}$, $\{e_i^j(\cdot)\}_{j=N_i}^{\infty}$ and $\{Z_i^j(\cdot)\}_{j=N_i}^{\infty}$ can be proven being relatively compact in $\mathcal{F}([0, T], \mathbb{P}^{n_{xi}})$, $\forall T > 0$. Then, the limits $P_i(\cdot)$ satisfy ODEs in (A.2). From the conclusion in Lemma 5, one has $\lim_{j\to\infty}P_{ij} = P_{i\infty}$ and $\lim_{j\to\infty}K_{ij} = K_{i\infty}$. Since $Q_i + K_{ij^*}^{\mathrm{T}}R_iK_{ij^*} - \gamma_i^2W_{ij^*}^{\mathrm{T}}W_{ij^*} > 0$, $A_i - B_iK_{ij^*} + D_iW_{ij^*}$ is Hurwitz. Then, one has $\lim_{j\to\infty}P_{ij} = P_i^*$ and $\lim_{j\to\infty}K_{ij} = K_i^*$. For pairs (X_i, U_i), (X_i, U_i) are obtained by solving static constrained optimization problems with the regulator equations in (9)–(10) as constraints and hence solve the regulator equations. The proof of Theorem 2 is completed. Besides, this theorem can be proven in an alternative way which does not rely on the asymptotic stability of the equilibrium point P_i^*. Clearly, there exists a sufficiently large integer $q_i^* > 0$ such that $\tilde{P}_{i(j+1)} \in B_q$, $\forall q > q_i^*, j > 0$ and hence $\{P_{ij}\}_{j=0}^{\infty}$ can be viewed as the Euler's method to approximate the ODE in (A.2). From Burden and Faires (2011, Section 5.2), one has that $\lim_{\sup\epsilon_k\to 0}\left\|P_{ij} - P_i\left(\sum_{k=0}^{j}\epsilon_k\right)\right\| = 0, \forall q > q_i^*$, which implies $\lim_{j\to\infty,\sup\epsilon_k\to 0}P_{ij} = P_i^*$. □

References

Arnold, W. F., & Laub, A. J. (1984). Generalized eigenproblem algorithms and software for algebraic Riccati equations. *Proceedings of the IEEE*, *72*(12), 1746–1754.

Başar, T., & Bernhard, P. (2008). *H_∞ optimal control and related minimax design problems*. Birkhäuser Boston.

Başar, T., & Olsder, G. J. (1998). *Dynamic noncooperative game theory*. SIAM.

Bian, T., & Jiang, Z.-P. (2015). Data-driven robust optimal control design for uncertain cascaded systems using value iteration. In *54th IEEE conference on decision and control* CDC, (pp. 7610–7615). IEEE.

Bian, T., & Jiang, Z.-P. (2016). Value iteration and adaptive dynamic programming for data-driven adaptive optimal control design. *Automatica*, *71*, 348–360.

Burden, R. L., & Faires, J. D. (2011). *Numerical analysis*. Cengage Learning.

Cai, H., Lewis, F. L., Hu, G., & Huang, J. (2017). The adaptive distributed observer approach to the cooperative output regulation of linear multi-agent systems. *Automatica*, *75*, 299–305.

Chen, B. M. (2000). *Robust and H_∞ control*. Springer.

Chen, C., Lewis, F. L., Xie, K., Xie, S., & Liu, Y. (2020). Off-policy learning for adaptive optimal output synchronization of heterogeneous multi-agent systems. *Automatica*, *119*, Article 109081.

Chen, C., Xie, L., Jiang, Y., Xie, K., & Xie, S. (2022). Robust output regulation and reinforcement learning-based output tracking design for unknown linear discrete-time systems. *IEEE Transactions on Automatic Control*, http://dx.doi.org/10.1109/TAC.2022.3172590, in press.

Ding, Z. (2017). Distributed adaptive consensus output regulation of network-connected heterogeneous unknown linear systems on directed graphs. *IEEE Transactions on Automatic Control*, *62*(9), 4683–4690.

Fan, J., Li, Z., Jiang, Y., Chai, T., & Lewis, F. L. (2018). Model-free linear discrete-time system H_∞ control using input-output data. In *2018 international conference on advanced mechatronic systems (ICAMechS)* (pp. 207–212). Zhengzhou, Henan, China.

Fan, J., Wu, Q., Jiang, Y., Chai, T., & Lewis, F. L. (2020). Model-free optimal output regulation for linear discrete-time lossy networked control systems. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *50*(11), 4033–4042.

Francis, B. A. (1977). The linear multivariable regulator problem. *SIAM Journal on Control and Optimization*, *15*(3), 486–505.

Gao, W., Deng, C., Jiang, Y., & Jiang, Z. (2022). Resilient reinforcement learning and robust output regulation under denial-of-service attacks. *Automatica*, *142*, Article 110366.

Gao, W., & Jiang, Z.-P. (2016). Adaptive dynamic programming and adaptive optimal output regulation of linear systems. *IEEE Transactions on Automatic Control*, *61*(12), 4164–4169.

Gao, W., Jiang, Y., & Davari, M. (2019). Data-driven cooperative output regulation of multi-agent systems via robust adaptive dynamic programming. *IEEE Transactions on Circuits and Systems II: Express Briefs*, *66*(3), 447–451.

Gao, W., Jiang, Z.-P., Lewis, F. L., & Wang, Y. B. (2018). Leader-to-formation stability of multiagent systems: An adaptive optimal control approach. *IEEE Transactions on Automatic Control*, *63*(10), 3581–3587.

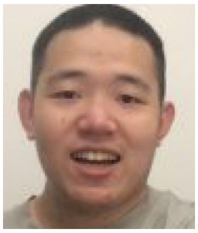
Gao, W., Mynuddin, M., Wunsch, D. C., & Jiang, Z.-P. (2021). Reinforcement learning-based cooperative optimal output regulation via distributed adaptive internal model. *IEEE Transactions on Neural Networks and Learning Systems*, http://dx.doi.org/10.1109/TNNLS.2021.3069728, in press.
Hu, W., Liu, L., & Feng, G. (2018). Cooperative output regulation of linear multi-agent systems by intermittent communication: A unified framework of time- and event-triggering strategies. *IEEE Transactions on Automatic Control*, *63*(2), 548–555.
Huang, J. (2004). *Nonlinear output regulation: Theory and applications*. SIAM.
Huang, J. (2017). The cooperative output regulation problem of discrete-time linear multi-agent systems by the adaptive distributed observer. *IEEE Transactions on Automatic Control*, *62*(4), 1979–1984.
Jiang, Y., Fan, J., & Chai, T. (2022). Data-driven optimal output regulation with assured convergence rate. *Acta Automatica Sinica*, *48*(4), 980–991.
Jiang, Y., Fan, J., Chai, T., & Lewis, F. L. (2019). Dual-rate operational optimal control for flotation industrial process with unknown operational model. *IEEE Transactions on Industrial Electronics*, *66*(6), 4587–4599.
Jiang, Y., Fan, J., Chai, T., Lewis, F. L., & Li, J. (2018). Tracking control for linear discrete-time networked control systems with unknown dynamics and dropout. *IEEE Transactions on Neural Networks and Learning Systems*, *29*(10), 4607–4620.
Jiang, Y., Fan, J., Chai, T., Li, J., & Lewis, F. L. (2018). Data-driven flotation industrial process operational optimal control based on reinforcement learning. *IEEE Transactions on Industrial Informatics*, *14*(5), 1974–1989.
Jiang, Y., Fan, J., Gao, W., Chai, T., & Lewis, F. L. (2020). Cooperative adaptive optimal output regulation of discrete-time nonlinear multi-agent systems. *Automatica*, *121*, Article 109149.
Jiang, Y., Gao, W., Na, J., Zhang, D., Hämäläinen, T. T., Stojanovic, V., & Lewis, F. L. (2022). Value iteration for adaptive optimal output regulation of linear continuous-time systems with assured convergence rate. *Control Engineering Practice*, *121*, Article 105042.
Jiang, Y., Kiumarsi, B., Fan, J., Chai, T., Li, J., & Lewis, F. L. (2020). Optimal output regulation of linear discrete-time systems with unknown dynamics using reinforcement learning. *IEEE Transactions on Cybernetics*, *50*(7), 3147–3156.
Jiang, Y., Zhang, K., Wu, J., Zhang, C., Xue, W., Chai, T., & Lewis, F. L. (2022). H_∞-based minimal energy adaptive control with preset convergence rate. *IEEE Transactions on Cybernetics*, *52*(10), 10078–10088.
Khalil, H. K. (1996). *Noninear systems*. New Jersey, USA: Prentice-Hall.
Kiumarsi, B., Lewis, F. L., Modares, H., Karimpour, A., & Naghibi-Sistani, M.-B. (2014). Reinforcement Q-learning for optimal tracking control of linear discrete-time systems with unknown dynamics. *Automatica*, *50*(4), 1167–1175.
Krener, A. J. (1992). The construction of optimal linear and nonlinear regulators. In *Systems, models and feedback: Theory and applications* (pp. 301–322). Springer.
Kučera, V. (1973). A review of the matrix Riccati equation. *Kybernetika*, *9*(1), 42–61.
Kushner, H. J., & Yin, G. G. (2003). *Stochastic approximation and recursive algorithms anda applications*. New York: Springer.
Lee, J. Y., Park, J. B., & Choi, Y. H. (2014). Integral reinforcement learning for continuous-time input-affine nonlinear systems with simultaneous invariant explorations. *IEEE Transactions on Neural Networks and Learning Systems*, *26*(5), 916–932.
Li, Z., Fan, J., Jiang, Y., & Chai, T. (2021). A model-free H_∞ method based on off-policy with output data feedback. *Acta Automatica Sinica*, *47*(9), 2182–2193.
Meng, Z. Y., Yang, T., Dimarogonas, D. V., & Johansson, K. H. (2015). Coordinated output regulation of heterogeneous linear systems under switching topologies. *Automatica*, *53*, 362–368.
Modares, H., Nageshrao, S. P., Lopes, G. A. D., Babuška, R., & Lewis, F. L. (2016). Optimal model-free output synchronization of heterogeneous systems using off-policy reinforcement learning. *Automatica*, *71*, 334–341.
Odekunle, A., Gao, W., Davari, M., & Jiang, Z.-P. (2020). Reinforcement learning and non-zero-sum game output regulation for multi-player linear uncertain systems. *Automatica*, *112*, Article 108672.
Pang, W., Fan, J., Jiang, Y., & Lewis, F. L. (2022). Optimal output regulation of partially linear discrete-time systems using reinforcement learning. *Acta Automatica Sinica*, *48*(9), 2242–2253.
Qin, J., Li, M., Shi, Y., Ma, Q., & Zheng, W. (2018). Optimal synchronization control of multiagent systems with input saturation via off-policy reinforcement learning. *IEEE Transactions on Neural Networks and Learning Systems*, *30*(1), 85–96.
Rizvi, S. A. A., & Lin, Z. (2018). Output feedback Q-learning for discrethe-time linear zero-sum games with application to the H-infinity control. *Automatica*, *95*, 213–221.
Rizvi, S. A. A., & Lin, Z. (2020). Output feedback adaptive dynamic programming for linear differential zero-sum games. *Automatica*, *122*, Article 109272.
Su, Y., & Huang, J. (2011). Cooperative output regulation of linear multi-agent systems. *IEEE Transactions on Automatic Control*, *57*(4), 1062–1066.
Vamvoudakis, K. G., & Lewis, F. L. (2010). Online actor–critic algorithm to solve the continuous-time infinite horizon optimal control problem. *Automatica*, *46*(5), 878–888.
Vamvoudakis, K. G., & Lewis, F. L. (2011). Multi-player non-zero-sum games: Online adaptive learning solution of coupled Hamilton-Jacobi equations. *Automatica*, *47*(8), 1556–1569.
Vamvoudakis, K. G., Lewis, F. L., & Hudas, G. R. (2012). Multi-agent differential graphical games: Online adaptive learning solution for synchronization with optimality. *Automatica*, *48*(8), 1598–1611.
van der Schaft, A. J. (1992). L_2-gain analysis of nonlinear systems and nonlinear state-feedback H_∞ control. *IEEE Transactions on Automatic Control*, *37*(6), 770–784.
Xue, W., Fan, J., Mejia, V. G. L., Jiang, Y., Chai, T., & Lewis, F. L. (2021). Off-policy reinforcement learning for tracking in continuous-time systems on two time-scales. *IEEE Transactions on Neural Networks and Learning Systems*, *32*(10), 4334–4346.
Xue, W., Fan, J. L., Mejia, V. G. L., Li, J., Jiang, Y., Chai, T., & Lewis, F. L. (2020). New methods for optimal operational control of industrial processes using reinforcement learning on two time-scales. *IEEE Transactions on Industrial Informatics*, *16*(5), 3085–3099.
Yan, Y., & Huang, J. (2016). Cooperative output regulation of discrete-time linear time-delay multi-agent systems. *IET Control Theory & Application*, *10*(16), 2019–2026.
Yang, Y., Modares, H., Wunsch, D. C., & Yin, Y. (2018). Leader-follower output synchronization of linear heterogeneous systems with active leader using reinforcement learning. *IEEE Transactions on Neural Networks and Learning Systems*, *29*(6), 2139–2153.
Yang, T., Saberi, A., Stoorvogel, A. A., & Grip, H. F. (2014). Output synchronization for heterogeneous networks of introspective right-invertible agents. *International Journal of Robust and Nonlinear Control*, *24*(13), 1821–1844.
Yuan, C. (2017). Robust H_∞ output regulation of uncertain linear fractional transformation systems with application to non-linear Chua's circuit. *IET Control Theory & Application*, *11*(12), 2012–2019.
Zhang, H., & Lewis, F. L. (2012). Adaptive cooperative tracking control of higher-order nonlinear systems with unknown dynamics. *Automatica*, *48*(7), 1432–1439.
Zhang, H., Liang, H., Wang, Z., & Feng, T. (2017). Optimal output regulation for heterogeneous multiagent systems via adaptive dynamic programming. *IEEE Transactions on Neural Networks and Learning Systems*, *28*(1), 18–29.

Yi Jiang was born in Ezhou, Hubei, China. He received the B.S. degree in automation, M.S. and Ph.D. degrees in control theory and control engineering from Information Science and Engineering College and State Key Laboratory of Synthetical Automation for Process Industries in Northeastern University, Shenyang, Liaoning, China in 2014, 2016 and 2020, respectively. From January to July, 2017, he was a Visiting Scholar with the UTA Research Institute, University of Texas at Arlington, TX, USA, and from March 2018 to March 2019, he was a Research Assistant with the University of Alberta, Edmonton, Canada. Currently, he is a Postdoc with the City University of Hong Kong, Hong Kong, China. His research interests include networked control systems, industrial process operational control, reinforcement learning and event-triggered control. Dr. Jiang is an Associate Editor for Advanced Control for Applications: Engineering and Industrial Systems, a Topic Editor for Frontiers in Control Engineering and the recipient of Excellent Doctoral Dissertation Award from Chinese Association of Automation (CAA) in 2021.

Weinan Gao received the B.Sc. degree in Automation from Northeastern University, Shenyang, China, in 2011, the M. Sc. degree in Control Theory and Control Engineering from Northeastern University, Shenyang, China, in 2013, and the Ph.D. degree in Electrical Engineering from New York University, Brooklyn, NY, USA in 2017. He is a Professor with the State Key Laboratory of Synthetical Automation for Process Industries at Northeastern University, Shenyang, China. Previously, he was an Assistant Professor of Mechanical and Civil Engineering at Florida Institute of Technology, Melbourne, FL, USA, an Assistant Professor of Electrical and Computer Engineering at Georgia Southern University, Statesboro, GA, USA, and a Visiting Professor of Mitsubishi Electric Research Laboratory (MERL), Cambridge, MA, USA. His research interests include reinforcement learning, adaptive dynamic programming (ADP), optimal control, cooperative adaptive cruise control (CACC), intelligent transportation systems, sampled-data control systems, and output regulation theory. He is the recipient of the best paper award in IEEE International Conference on Real-time Computing and Robotics (RCAR) in 2018,

and the David Goodman Research Award at New York University in 2019.

Dr. Gao is an Associate Editor of IEEE Transactions on Neural Networks and Learning Systems, IEEE/CAA Journal of Automatica Sinica, Control Engineering Practice, Neurocomputing and IEEE Transactions on Circuits and Systems II: Express Briefs, a member of Editorial Board of Neural Computing and Applications, and a Technical Committee member in IEEE Control Systems Society on Nonlinear Systems and Control and in IFAC TC 1.2 Adaptive and Learning Systems.



Jin Wu was born in May 1994, in Zhenjiang, China. He received the B.S. degree from the University of Electronic Science and Technology of China, Chengdu, China. He is currently working toward the Ph.D. degree with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong. Since 2018, he has been a Research Assistant with the Hong Kong University of Science and Technology. From 2012 to 2018, he was in the UAV industry. In 2013, he was a visiting student with Groep T, KU Leuven. From 2019 to 2020, he was with Tencent Robotics X, Shenzhen, China.

Tianyou Chai received the Ph.D. degree in control theory and engineering from Northeastern University, Shenyang, China, in 1985. He has been with the Research Center of Automation, Northeastern University, Shenyang, China, since 1985, where he became a Professor in 1988 and a Chair Professor in 2004. He is the founder and Director of the Center of Automation, which became a National Engineering and Technology Research Center in 1997. His current research interests include adaptive control, intelligent decoupling control, integrated plant control and systems, and the development of control technologies with applications to various industrial processes.

Prof. Chai is a member of the Chinese Academy of Engineering, an academician of International Eurasian Academy of Sciences, and an IFAC Fellow. He is a Distinguished Visiting Fellow of The Royal Academy of Engineering (UK) and an Invitational Fellow of Japan Society for the Promotion of Science (JSPS). For his contributions, he has won three prestigious awards of National Science and Technology Progress, the 2002 Technological Science Progress Award from the Ho Leung Ho Lee Foundation, the 2007 Industry Award for Excellence in Transitional Control Research from the IEEE Control Systems Society, and the 2010 Yang Jia-Chi Science and Technology Award from the Chinese Association of Automation.

Frank L. Lewis received the bachelor's degree in physics/electrical engineering and the M.S.E.E. degree from Rice University, Houston, TX, USA, the M.S. degree in aeronautical engineering from the University of West Florida, Pensacola, FL, USA, and the Ph.D. degree from the Georgia Institute of Technology, Atlanta, GA, USA. He is currently a Distinguished Scholar Professor and Distinguished Teaching Professor with the University of Texas at Arlington, Fort Worth, TX, USA, the Moncrief-O'Donnell Chair of the University of Texas at Arlington Research Institute, Fort Worth. He is involved in feedback control, intelligent systems, cooperative control systems, and nonlinear systems.

Prof. Lewis is a member of the National Academy of Inventors, a fellow of the International Federation of Automatic Control and the Institute of Measurement and Control, U.K., a Texas Board of Professional Engineer, a U.K. Chartered Engineer, and a founding member of the Board of Governors of the Mediterranean Control Association. He was a recipient of the Fulbright Research Award, the NSF Research Initiation Grant, the Terman Award from the American Society for Engineering Education, the Gabor Award from the International Neural Network Society, the Honeywell Field Engineering Medal from the Institute of Measurement and Control, U.K., the Neural Networks Pioneer Award from IEEE Computational Intelligence Society, the Outstanding Service Award from the Dallas IEEE Section, and the Texas Regents Outstanding Teaching Award in 2013. He was elected as an Engineer of the year by the Fort Worth IEEE Section. He was listed in the Fort Worth Business Press Top 200 Leaders in Manufacturing.