Differential graphical games for H-infinity control of linear heterogeneous multiagent systems

Farnaz Adib Yaghmaie, Kristian Hengster Movric, Frank L. Lewis and Rong Su

INTERNATIONAL JOURNAL OF ROBUST AND NONLINEAR CONTROL
Int. J. Robust. Nonlinear Control 2019; **00**:1–21
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/rnc

Differential Graphical Games for H_∞ Control of Linear Heterogeneous Multi-agent Systems

Farnaz Adib Yaghmaie [1 4], Kristian Hengster Movric [2], Frank L. Lewis[3], Rong Su [4]

[1] *Department of Electrical Engineering, Linköping University, 58183 Linköping, Sweden,*
[2] *Department of Control Engineering, Faculty of Electrical Engineering, Czech Technical University in Prague, Prague, Czech,*
[3] *University of Texas at Arlington Research Institute, The University of Texas at Arlington, Texas, USA and Qian Ren Consulting Professor, Northeastern University, Shenyang 110036, China,*
[4] *School of Electrical & Electronic Engineering, Nanyang Technological University, Singapore.*

SUMMARY

Differential graphical games have been introduced in the literature to solve state synchronization problem for linear homogeneous agents. When the agents are heterogeneous, the previous notion of graphical games cannot be used anymore and a new definition is required. In this paper, we define a novel concept of differential graphical games for linear heterogeneous agents subject to external unmodeled disturbances which contain the previously introduced graphical game for homogeneous agents as a special case. Using our new formulation, we can solve both the output regulation and H_∞ output regulation problems. Our graphical game framework yields coupled Hamilton-Jacobi-Bellman equations which are, in general, impossible to solve analytically. Therefore, we propose a new actor-critic algorithm to solve these coupled equations numerically in real time. Moreover, we find an explicit upper bound for the overall $\mathcal{L}_2$-gain of the output synchronization error with respect to disturbance. We demonstrate our developments by a simulation example. Copyright © 2019 John Wiley & Sons, Ltd.

Received ...

1. INTRODUCTION

The next generation of networked systems is currently emerging from a number of different engineering domains. We already have examples of Internet of Things (IoT), Industry 4.0, smart cities and various other cyber-physical systems characterized by the requirement to coordinate efforts across large networks of different agents. In the near future, these systems are expected to increase in importance, with an ever-widening range of applications. Main design goals there are versatility, flexibility, easy real time reconfigurability, low communication burden, along with robustness to component failures and resilience to disturbances. These complex interconnected multi-agent systems create the need for novel team decision, distributed control, optimization and online computation methodologies.

A multi-agent system is defined as a group of interconnected dynamical systems interacting to achieve a desired collective behavior like state synchronization [1, 2, 3], output regulation [4, 5, 6], cluster synchronization [7], formation control [8], [9], etc. The canonical distributed control problem of state synchronization is usually defined for homogeneous agents where it is possible to use the local state synchronization error [1, 2, 3]. If the agents are *heterogeneous*, meaning that they

*Correspondence to: Farnaz Adib Yaghmaie, Department of Electrical Engineering, Linköping University, 58183 Linköping, Sweden. Email: farnaz.adib.yaghmaie@liu.se. Phone: +46-762909978.

have different internal dynamics, it is not generally meaningful to achieve state synchronization; instead, one may consider an output regulation problem where a distributed controller is designed such that the outputs of all agents synchronize to a reference trajectory while the effect of *modeled disturbances*, i.e. disturbance with known dynamic model, is rejected. In these cases, the dynamics of the reference trajectory and the disturbance model are usually combined into a single dynamic model named *exo-system* in the literature [10, 4]. One possible solution to such output regulation problems is by the means of *Internal Model Principle (IMP)* where the idea is to incorporate an internal model of the exo-system in the dynamic controller of each agent [10, 11, 12, 4, 6]. If the agents are additionally subject to *unmodeled disturbances* then the IMP alone cannot be used for disturbance rejection and H_∞ control methods are required. A significant part of the research on H_∞ control of multi-agent systems is on H_∞ state synchronization of homogeneous agents [13, 14] and more recently H_∞ output regulation of heterogeneous agents [15, 16, 17].

In the context of multi-agent systems, various optimization methods are utilized to achieve the required design characteristics. One has the conventional centralized optimization [18, 19] as well as more recent distributed approaches [20, 21] aiming at flexibility, reconfigurability, and robustness. The optimal control of a single dynamic system, also called a *single-player game* [22, 23], is the simplest dynamic optimization problem. However, this approach is usually found lacking in robustness to external disturbances acting on the system and possibly subsystem failures. The optimal control of a dynamic system subject to unmodeled disturbance is termed *two-player zero-sum game* where the control player tries to minimize a cost function while the disturbance player maximizes it [24, 25]. More recently, the idea of *graphical game* is introduced to capture and exploit the locality and influences of dynamic systems on each other. In graphical games, the dynamics and the objective function of each player are influenced by other players in the neighborhood. A graphical game for *homogeneous agents* i.e. agents with the identical internal dynamics is suggested in [20] to achieve state synchronization using local state synchronization error. This concept is extended to H_∞ graphical game where the agents are subject to external disturbances [21]. A graphical game for cluster synchronization of homogeneous agents is given in [7] where the agents in each cluster are synchronized to the reference trajectory in that cluster while different clusters have different reference trajectories. These graphical games provide a suitable platform for distributed optimal controller designs for the identical agents. The choice of individual player objective function, as discussed above, ensures that the multi-agent system achieves the desired collective behavior e.g. state, H_∞-state or cluster synchronization [20, 21], [7].

Reinforcement Learning (RL) is concerned with learning optimal policies from interaction with an environment [26]. In RL, a decision-making system modifies its control policy based on the stimuli received in response to its previous policies to optimize the cost. In this sense, RL implies a cause and effect relationship between policies and costs [27], and as such, RL based frameworks enjoy optimality and adaptivity. Over the past few years, dynamic programming is utilized to develop RL techniques for adaptive-optimal control of dynamic systems, see [22, 24, 28, 25, 19, 20, 21] to name a few. Hence, RL is by now fairly standard in solving the single-player optimal and robust control problems [22, 23]. However, as of recently, RL can also be used to tackle even more complicated optimization structures like multiple interconnected dynamics system, namely a multi-player game, where a single centralized dynamics contains the dynamics of all players [18, 19]. Likewise, RL methods have also been successfully applied in graphical games [20, 21], [7]. Moreover, faced with complicated coupled design equations arising from such problems, RL remains the only viable method applicable in real time. Such recent approaches are particularly appropriate for problems arising from the domain of multi-agent systems.

In this paper, we aim to develop a novel graphical game framework for linear heterogeneous multi-agent systems. For this purpose, we bring together distributed control of heterogeneous multi-agent systems, graphical games, and reinforcement learning techniques. There are four main contributions in this paper. (1) We define the novel concept of graphical games for heterogeneous agents as opposed to homogeneous agents considered in [20, 21]. This allows us to achieve output regulation among heterogeneous agents. Graphical game for heterogeneous agents is also considered in [12], however, the communication graph in [12] is required to be acyclic (i.e. there is no-loop in the graph). This restrictive assumption significantly simplifies the formulation and decouples the controller design of each agent from the others. (2) We assume that the agents are subject to unmodeled disturbances and define an H_∞ graphical game for heterogeneous agents. The only reference pertaining to H_∞ control of multi-agent systems in the graphical game framework [21] considers only homogeneous agents. (3) H_∞ graphical game for heterogeneous agents results in coupled Hamilton-Jacobi-Bellman equations that are difficult to solve analytically. We use RL and develop new actor-critic networks to obtain solutions to these equations. In contrast, the actor-critic networks in [20, 21] can be used only for homogeneous agents. (4) We obtain an upper bound for the $\mathcal{L}_2$-gain of output synchronization error with respect to unmodeled disturbances; in contrast to [21], which does not calculate the upper bound but only contends its existence.

The rest of the paper is organized as follows. In Section 2, we discuss notation conventions and review preliminaries. In Section 3, we define the H_∞ output regulation problem. In Section 4, we define the novel graphical games for linear heterogeneous agents. In Section 5, we present our results regarding H_∞ output regulation using graphical games and derive the resulting overall $\mathcal{L}_2$-gain of output synchronization error with respect to disturbances. In Section 6, we suggest a distributed online procedure for solving the H_∞ output regulation graphical game. We demonstrate the validity of theoretical developments with a simulation example in Section 7. We conclude the paper in Section 8.

2. PRELIMINARIES

2.1. Notation

The following notation will be used throughout this paper. Let $\mathbb{R}^{n\times m}$ be the set of $n \times m$ real matrices. I_n denotes the identity matrix of dimension $n \times n$ and $\mathbf{1}_N$ is an N column vector of 1. $\mathbf{0}$ denotes a matrix of zeros with compatible dimensions. The Kronecker product of two matrices A and B is denoted by $A \otimes B$. The positive (semi) definiteness constraint on the matrix P is expressed as $P > 0$ ($P \geq 0$). Let $A_i \in \mathbb{R}^{n_i \times m_i}$ for $i = 1, ..., N$, where N is a positive integer. The operator $\underset{1:N}{Diag}\{A_i\}$ is defined as

$$\underset{1:N}{Diag}\{A_i\} = \begin{bmatrix} A_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & A_2 & \dots & \mathbf{0} \\ \vdots & \mathbf{0} & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & A_N \end{bmatrix}. \tag{1}$$

The maximum singular value of a matrix A is denoted by $\overline{\sigma}(A)$ and its kernel is denoted by $Ker(A)$. An eigenvalue of a square matrix A is denoted by $\lambda_i(A)$.

2.2. $\mathcal{L}_p$-norm

For $p \in [1, +\infty)$, let $\mathcal{L}_p = \mathcal{L}_p^n[0, +\infty)$ denote the space of functions $a(t) \in \mathbb{R}^n$ such that $t \to |a(t)|^p$ is integrable over $[0, +\infty)$, where $|a(t)|$ is the instantaneous Euclidean norm of the vector $a(t)$. The $\mathcal{L}_p$-norm of $a(t) \in \mathcal{L}_p^n[0, \infty)$ is defined as

$$\|a(t)\|_{\mathcal{L}_p} = \left(\int_0^\infty |a(\tau)|^p d\tau\right)^{1/p} < +\infty.$$

2.3. Graph theory

Suppose that the interaction among the followers is represented by an undirected graph $\mathcal{G} = (V, \mathcal{E})$ with a finite set of N nodes $V = \{v_1, ..., v_N\}$ and a set of undirected edges $\mathcal{E} \subseteq V \times V$. $E = [\alpha_{ij}]$ is the *adjacency matrix* with $\alpha_{ij} = 1$ if $(v_j, v_i) \in \mathcal{E}$ and $\alpha_{ij} = 0$ otherwise. Since the graph is undirected, the adjacency matrix is symmetric. The graph is simple, *i.e.* $\alpha_{ii} = 0,\ i = 1, ..., N$. A

DOI: 10.1002/rnc

path from node v_i to node v_j is a sequence of edges joining v_i to v_j. If there exists a path from node v_i to node v_j then it is said that node v_i is *reachable* from v_j. The set of *neighbors of node* v_i is $N_i = \{v_j : (v_j, v_i) \in \mathcal{E}\}$. For graph $\mathcal{G}$, the *in-degree matrix* D is a diagonal matrix $D = \underset{1:N}{Diag}\{d_i\}$ with $d_i = \sum_{j \in N_i} \alpha_{ij}$. The *laplacian matrix* for graph $\mathcal{G}$ is defined as $L = D - E$. Suppose that matrix $G = \underset{1:N}{Diag}\{g_i\}$ shows the pinning from the leader (v_0) to the followers ($v_1, ..., v_N$). Then, $g_i = 1$ if there is a link between the leader and follower i and $g_i = 0$ otherwise. Denote the augmented graph by $\bar{\mathcal{G}} = (\bar{V}, \bar{\mathcal{E}})$ which is obtained by attaching node v_0 and its outgoing edges to $\mathcal{G}$. Graph $\bar{\mathcal{G}}$ shows the interaction among the followers and the leader.

A graph $\mathcal{G} = (V, \mathcal{E})$ is *connected* if there exists a path from v_i to v_j for all $v_i, v_j \in V$. If the initial and the terminal nodes of a path are the same, the path is called a *cycle*. A graph without any cycle is named an *acyclic graph*. A graph is a *connected tree* if every node except one node, called *the root*, has in-degree equal to one. The root node has its in-degree equal to zero. A graph has a *spanning tree* if there exists a tree containing every node in V.

3. H_∞ OUTPUT REGULATION PROBLEM

Consider a set of $N + 1$ heterogeneous agents with N followers given as LTI systems

$$\dot{x}_i = A_i x_i + B_i u_i + P_i \omega_i, \tag{2}$$

$$y_i = C_i x_i + E_i u_i, \tag{3}$$

$$z_i = D_i x_i, \tag{4}$$

and a leader given by

$$\dot{\xi}_0 = X \xi_0, \tag{5}$$

$$y_0 = R_1 \xi_0, \tag{6}$$

$$z_0 = R_2 \xi_0, \tag{7}$$

in which $x_i \in \mathbb{R}^{n_i}$, $y_i \in \mathbb{R}^p$, $z_i \in \mathbb{R}^q$ and $u_i \in \mathbb{R}^{m_i}$ denote the state, the synchronization output, the measured output and the control signal for follower $i = 1, ..., N$. The $\xi_0 \in \mathbb{R}^l$, $y_0 \in \mathbb{R}^p$ and $z_0 \in \mathbb{R}^q$ denote the state, the synchronization output and the measured output of the leader. All followers are subject to external unmodeled disturbances $\omega_i \in \mathcal{L}_2$. The motivation behind introducing two outputs y_i, z_i is to achieve synchronization in outputs y_i, by communicating the measured outputs z_i.

We suggest a distributed static output-feedback controller of the following form

$$u_i = K_i e_{zi}, \tag{8}$$

$$e_{zi} = \sum_{j \in N_i} \alpha_{ij}(z_i - z_j) + g_i(z_i - z_0), \tag{9}$$

where e_{zi} is the local neighborhood error in z-outputs and $K_i \in \mathbb{R}^{m_i \times q}$. Define the y-output synchronization error and the z-output synchronization error as

$$\delta_{yi} = y_i - y_0, \tag{10}$$

$$\delta_{zi} = z_i - z_0. \tag{11}$$

Let $x = [x_1^T, ..., x_N^T]^T$, $y = [y_1^T, ..., y_N^T]^T$, $z = [z_1^T, ..., z_N^T]^T$, $\omega = [\omega_1^T, ..., \omega_N^T]^T$, $\delta_y = [\delta_{y1}^T, ..., \delta_{yN}^T]^T$ and $\delta_z = [\delta_{z1}^T, ..., \delta_{zN}^T]^T$ denote the overall vectors of x_i, y_i, z_i, ω_i, δ_{yi} and δ_{zi} respectively. Then, the overall closed-loop system of all followers, their controllers and the leader is given by the following

$$\begin{aligned} \dot{x} &= A_{cl} x + B_{cl} \xi_0 + P_{cl} \omega, \\ \dot{\xi}_0 &= X \xi_0, \\ \delta_y &= C_{cl} x - E_{cl} \xi_0, \\ \delta_z &= D_{cl} x - \mathbf{1}_N \otimes R_2 \xi_0, \end{aligned} \tag{12}$$

where

$$\begin{aligned}
A_{cl} &= \underset{1:N}{Diag}\{A_i\} + \underset{1:N}{Diag}\{B_iK_i\}((L+G)\otimes I_q)\underset{1:N}{Diag}\{D_i\},\\
B_{cl} &= -\underset{1:N}{Diag}\{B_iK_i\}((L+G)\otimes I_q)(\mathbf{1}_N \otimes R_2),\\
C_{cl} &= \underset{1:N}{Diag}\{C_i\} + \underset{1:N}{Diag}\{E_iK_i\}((L+G)\otimes I_q)\underset{1:N}{Diag}\{D_i\},\\
E_{cl} &= \underset{1:N}{Diag}\{E_iK_i\}((L+G)\otimes I_q)(\mathbf{1}_N \otimes R_2) + \mathbf{1}_N \otimes R_1,\\
D_{cl} &= \underset{1:N}{Diag}\{D_i\},\ P_{cl} = \underset{1:N}{Diag}\{P_i\}.
\end{aligned}$$

Now, we define the H_∞ output regulation problem.

Problem 1 (H_∞ output regulation problem for linear heterogeneous multi-agent systems)
Consider a group of $N+1$ heterogeneous LTI systems defined by (2-7). Design the feedback gains K_i such that for $\forall x_i(0),\ i=1,...,N$, the closed-loop system of (2)-(7) using (8) achieves the following properties:

a. For $\omega \equiv \mathbf{0}$ and $\xi_0 \equiv \mathbf{0}$, the origin of the system $\dot{x} = A_{cl}x$ is asymptotically stable.
b. For $\omega \equiv \mathbf{0}$, we have $\delta_y \to \mathbf{0}$ and $\delta_z \to \mathbf{0}$ as $t \to +\infty$ for all initial conditions $\xi_0(0)$.
c. For $\omega \in \mathcal{L}_2$ and $T > 0$

$$L^2_{\delta_y\omega} = \frac{\int_0^T \|\delta_y\|_2^2 d\tau}{\int_0^T \|\omega\|_2^2 d\tau} < +\infty. \tag{13}$$

Properties a.-b. define the output regulation problem [29, 30] in absence of unmodeled disturbance; i.e. $\omega \equiv \mathbf{0}$. Property c. concerns H_∞ control in presence of unmodeled disturbances [31]. We make the following assumptions throughout this paper.

Assumption 1
The augmented graph $\bar{\mathcal{G}} = (\bar{V}, \bar{\mathcal{E}})$ contains a spanning tree with the leader as its root node.

Assumption 2
The leader's dynamics X in (5) does not have any strictly stable pole.

Assumption 3
The triple (A_i, B_i, D_i) is output-feedback stabilizable.

According to Lemma 1.4 of [30], the closed-loop system (12) achieves y- and z-outputs synchronization to the outputs of the leader for $\omega \equiv \mathbf{0}$ only if there exists an invariant subspace for the closed-loop system (12), where $y_i = y_0$ and $z_i = z_0$ for $i = 1, ..., N$

$$\begin{aligned}
\begin{bmatrix} A_{cl} & B_{cl} \\ \mathbf{0} & X \end{bmatrix} \begin{bmatrix} \Pi \\ I_l \end{bmatrix} &= \begin{bmatrix} \Pi \\ I_l \end{bmatrix} X,\\
\begin{bmatrix} C_{cl} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \Pi \\ I_l \end{bmatrix} &= E_{cl},\\
\begin{bmatrix} D_{cl} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \Pi \\ I_l \end{bmatrix} &= \mathbf{1}_N \otimes R_2,
\end{aligned} \tag{14}$$

where $\Pi \in \mathbb{R}^{\sum_{i=1}^N n_i \times l}$. Using the last equation in the first and second equations, (14) is simplified to

$$\begin{aligned}
\underset{1:N}{Diag}\{A_i\}\Pi &= \Pi X,\\
\underset{1:N}{Diag}\{C_i\}\Pi &= \mathbf{1}_N \otimes R_1,\\
\underset{1:N}{Diag}\{D_i\}\Pi &= \mathbf{1}_N \otimes R_2.
\end{aligned}$$

Let $\Pi = \begin{bmatrix}\Pi_1^T & \Pi_2^T & \dots & \Pi_N^T\end{bmatrix}^T$, where $\Pi_i \in \mathbb{R}^{n_i \times l}$ has full column rank. Then, the aforementioned necessary condition for y- and z-outputs synchronization is summarized in the following assumption.

Assumption 4
For each $i = 1, ..., N$, there exists a matrix $\Pi_i \in \mathbb{R}^{n_i \times l}$ of full column rank such that

$$\begin{aligned} A_i\Pi_i &= \Pi_i X, \\ C_i\Pi_i &= R_1, \\ D_i\Pi_i &= R_2. \end{aligned} \tag{15}$$

Remark 1
We make Assumption 2 without loss of generality (remark 1.3 in [30]). A leader not satisfying Assumption 2 is asymptotically identical to a simpler leader, one of a smaller order, that satisfies it. As we are mainly interested in the long-term behavior, we can restrict our attention to leaders satisfying Assumption 2. Any difference would only be in the transient, with no long-term effects. In fact, if the output regulation problem is solvable by any controller under Assumption 2, then it is also solvable by the same controller even if this Assumption is violated [30].

3.1. Coordinate Transformations

In this subsection, we introduce a coordinate transformation which is useful in formulating the output regulation problem and the graphical game in Section 4. Building on Assumption 4, supplement the columns of Π_i in (15) by a set of linearly independent columns of $\Psi_i \in \mathbb{R}^{n_i \times (n_i - l)}$ to form a complete basis $T_i = [\Pi_i \ \Psi_i] \in \mathbb{R}^{n_i \times n_i}$ of the single-agent state space $\mathbb{R}^{n_i}$. Then in such basis, one has the transformed state $[\xi_i^T, \nu_i^T]^T$

$$x_i = \begin{bmatrix}\Pi_i & \Psi_i\end{bmatrix}\begin{bmatrix}\xi_i \\ \nu_i\end{bmatrix} = T_i\begin{bmatrix}\xi_i \\ \nu_i\end{bmatrix}. \tag{16}$$

Define the following matrices

$$\hat{A}_i := T_i^{-1}A_iT_i = \begin{bmatrix}X & F_i \\ \mathbf{0} & M_i\end{bmatrix}, \quad \hat{B}_i := T_i^{-1}B_i = \begin{bmatrix}\hat{B}_{1i} \\ \hat{B}_{2i}\end{bmatrix}, \quad \hat{P}_i := T_i^{-1}P_i = \begin{bmatrix}\hat{P}_{1i} \\ \hat{P}_{2i}\end{bmatrix}. \tag{17}$$

Define the local neighborhood error in ξ_is

$$e_{\xi_i} = \sum_{j=1}^{N} \alpha_{ij}(\xi_i - \xi_j) + g_i(\xi_i - \xi_0). \tag{18}$$

Let $\epsilon_i = [e_{\xi_i}^T, \nu_i^T]^T$. Then, the dynamics of system (2) and the control (8) in the transformed coordinates ϵ_i read

$$\dot{\epsilon}_i = \bar{A}_i\epsilon_i + \bar{B}_iu_i + \bar{P}_i\omega_i - \sum_{j=1}^{N}\alpha_{ij}(\underline{B}_ju_j + \underline{P}_j\omega_j + \underline{F}_j\nu_j), \tag{19}$$

$$u_i = K_iR_2e_{\xi_i} + (d_i + g_i)K_iD_i\Psi_i\nu_i - K_i\sum_{j=1}^{N}\alpha_{ij}D_j\Psi_j\nu_j, \tag{20}$$

where

$$\begin{aligned} &\bar{A}_i := \begin{bmatrix}X & (d_i+g_i)F_i \\ \mathbf{0} & M_i\end{bmatrix}, \quad \bar{B}_i := \begin{bmatrix}(d_i+g_i)\hat{B}_{1i} \\ \hat{B}_{2i}\end{bmatrix}, \quad \bar{P}_i := \begin{bmatrix}(d_i+g_i)\hat{P}_{1i} \\ \hat{P}_{2i}\end{bmatrix}, \\ &\underline{B}_j := \begin{bmatrix}\hat{B}_{1j} \\ \mathbf{0}\end{bmatrix}, \quad \underline{P}_j := \begin{bmatrix}\hat{P}_{1j} \\ \mathbf{0}\end{bmatrix}, \quad \underline{F}_j := \begin{bmatrix}F_j \\ \mathbf{0}\end{bmatrix}. \end{aligned} \tag{21}$$

Equation (19) describes agent i dynamics in the new $\epsilon_i^T = [e_{\xi_i}^T, \nu_i^T]$-coordinates. The following technical lemma can be used to simplify the dynamics in (20).

Lemma 1
One can always select Ψ_i in (16) such that $F_i = \mathbf{0}$ in (17).

Proof
From the definition of $\hat{A}_i$ in (17)

$$A_i \begin{bmatrix}\Pi_i & \Psi_i\end{bmatrix} = \begin{bmatrix}\Pi_i & \Psi_i\end{bmatrix} \begin{bmatrix} X & F \\ \mathbf{0} & M_i \end{bmatrix}.$$

Hence, one has $A_i\Psi_i - \Psi_i M_i = \Pi_i F_i$, which is a Sylvester equation. Using the Kronecker product's property, this is equivalent to $(I_{\tilde{n}_i} \otimes A_i - M_i^T \otimes I_{n_i})vec(\Psi_i) = vec(\Pi_i F_i)$, where $\tilde{n}_i = n_i - l$. If one wants $F_i = \mathbf{0}$, then one should select

$$vec(\Psi_i) \in null(I_{\tilde{n}_i} \otimes A_i - M_i^T \otimes I_{n_i}). \tag{22}$$

Let $\lambda_k \in spec(A_i)$, $\tilde{\lambda}_k \in spec(M_i)$. Then, $(\lambda_k - \tilde{\lambda}_k) \in spec(I_{\tilde{n}_i} \otimes A_i - M_i^T \otimes I_{n_i})$. Since $spec(M_i) \subset spec(A_i)$, we have $0 \in spec(I_{\tilde{n}_i} \otimes A_i - M_i^T \otimes I_{n_i})$ and it is always possible to satisfy (22). For a more general result, please see Theorem 4.4.14 of [32]. □

Using Lemma 1, we hereafter assume that we have selected Ψ_i such that $F_i = \mathbf{0}$ for $i = 1, ..., N$.

3.2. Closed-loop system and node error dynamics for graphical game

In this subsection, we define the node error dynamics which are suitable for graphical games formulation for heterogeneous agents fully developed in Section 4. To do so, we express the closed-loop system of (19) using the control (20)

$$\begin{aligned}
\dot{e}_{\xi_i} =& X e_{\xi_i} + (d_i+g_i)\tilde{B}_{1i}K_iR_2e_{\xi_i} + (d_i+g_i)^2\tilde{B}_{1i}K_iD_i\Psi_i\nu_i + (d_i+g_i)\tilde{P}_{1i}\omega_i \\
&- (d_i+g_i)\tilde{B}_{1i}K_i\sum_{j=1}^N \alpha_{ij}D_j\Psi_j\nu_j - \sum_{j=1}^N \alpha_{ij}\tilde{B}_{1j}K_jR_2e_{\xi_j} \\
&- \sum_{j=1}^N \alpha_{ij}(d_j+g_j)\tilde{B}_{1j}K_jD_j\Psi_j\nu_j + \sum_{j=1}^N \alpha_{ij}\tilde{B}_{1j}K_j\sum_{l=1}^N \alpha_{jl}D_l\Psi_l\nu_l - \sum_{j=1}^N \alpha_{ij}\tilde{P}_{1j}\omega_j, \\
\dot{\nu}_i =& \tilde{A}_i\nu_i + \tilde{B}_{2i}K_iR_2e_{\xi_i} + (d_i+g_i)\tilde{B}_{2i}K_iD_i\Psi_i\nu_i + \tilde{P}_{2i}\omega_i - \tilde{B}_{2i}K_i\sum_{j=1}^N \alpha_{ij}D_j\Psi_j\nu_j
\end{aligned} \tag{23}$$

Define

$$\beta_i := -K_i\sum_{j=1}^N \alpha_{ij}D_j\Psi_j\nu_j + (d_i+g_i)(K_iD_i\Psi_i - \bar{K}_i)\nu_i,\ \beta = \begin{bmatrix}\beta_1^T, & ..., & \beta_N^T\end{bmatrix}^T, \tag{24}$$

$$u_{op_i} := K_iR_2e_{\xi_i} + (d_i+g_i)\bar{K}_i\nu_i, \tag{25}$$

so that $u_i = u_{op_i} + \beta_i$. In (24)-(25), $\bar{K}_i \in \mathbb{R}^{m_i\times(n_i-l)}$ is a gain matrix which has no effect on the design of the controller but we introduce it to facilitate the following theoretical developments of the optimal design. The signal β_i in (24) contains the term $-(d_i+g_i)\bar{K}_i\nu_i$ and u_{op_i} in (25) contains $(d_i+g_i)\bar{K}_i\nu_i$. Clearly, this term has no effect on the design of the controller because u_{op_i} and β_i always appear added together.

By definitions in (24)-(25), the closed-loop system (23) can be represented as

$$\dot{\epsilon}_i = \bar{A}_i\epsilon_i + \bar{B}_iu_{op_i} + \bar{P}_i\omega_i + \bar{B}_i\beta_i - \sum_{j=1}^N \alpha_{ij}(\underline{B}_ju_{op_j} + \underline{P}_j\omega_j + \underline{B}_j\beta_j). \tag{26}$$

DOI: 10.1002/rnc

We call (26) *the node error dynamics for graphical games for heterogeneous agents*. It is important to point out that (26) is in a standard form of the dynamics usually considered in the graphical game framework [21, 20], *i.e.* it depends on the states, policies and disturbances of other agents in the neighborhood. Note however that while the agents in [21, 20] are homogeneous, the dynamics in (26) are heterogeneous.

Remark 2
Note that ν_is and the related β_is stem from agents heterogeneity and they pose additional complications for control design that do not arise in identical agents. In this more general setup with heterogeneous agents, we develop the graphical game in Section 4 in the sequel, whereas the existing results along similar lines consider only identical agents. For a special case of identical agents, our formulation indeed reduces to the familiar one in [21, 20].

In Section 4, we use (26) to develop the graphical game and devise the appropriate controls u_{op_i}. In Section 5 then, we bring additional conditions guaranteeing that the results of Section 4 solve the original Problem 1.

4. GRAPHICAL GAME FOR LINEAR HETEROGENEOUS AGENTS

In this section, we use the machinery of graphical games and define an H_∞ graphical game for linear heterogeneous agents. Using this graphical game framework, we ultimately solve our H_∞ output regulation problem of Section 3.

4.1. H_∞ graphical game formulation

As it can be seen from (26), the dynamics of each agent depends on the neighboring agents. In a similar way, we define the $\mathcal{L}_2$-condition for agent i such that it contains the policies and disturbances of other agents in its neighborhood. Let $u_{op_{-i}} = \{u_{op_j},\ j \in N_i\}$, $\omega_{-i} = \{\omega_j,\ j \in N_i\}$ and $\beta_{-i} = \{\beta_j,\ j \in N_i\}$. Define the $\mathcal{L}_2$-condition for the dynamics (26) as

$$\begin{aligned}&\int_0^T \{\epsilon_i^T Q_i \epsilon_i + u_{op_i}^T R_{ii} u_{op_i} + \sum_{j=1}^N \alpha_{ij} u_{op_j}^T R_{ij} u_{op_j}\} d\tau \leq \\ &\quad \gamma_{\xi\nu_i}^2 \int_0^T \{\omega_i^T S_{1ii}\omega_i + \sum_{j=1}^N \alpha_{ij}\omega_j^T S_{1ij}\omega_j\} d\tau + \gamma_{\beta_i}^2 \int_0^T \{\beta_i^T S_{2ii}\beta_i + \sum_{j=1}^N \alpha_{ij}\beta_j^T S_{2ij}\beta_j\} d\tau,\end{aligned} \tag{27}$$

where $Q_i > 0,\ R_{ii} > 0,\ R_{ij} \geq 0,\ S_{1ii} > 0,\ S_{1ij} \geq 0,\ S_{2ii} > 0,\ S_{2ij} \geq 0,\ \gamma_{\xi\nu_i} > 0,\ \gamma_{\beta_i} > 0$ and $T > 0$. The $\mathcal{L}_2$-condition in (27) is equivalent to the optimization of the following quadratic performance index

$$\begin{aligned} J_i(\epsilon_i(0), u_{op_i}, u_{op_{-i}}, \omega_i, \omega_{-i}, \beta_i, \beta_{-i}) &= \int_0^{+\infty} L_i(\epsilon_i, u_{op_i}, u_{op_{-i}}, \omega_i, \omega_{-i}, \beta_i, \beta_{-i}) d\tau \\ &= \int_0^{+\infty} \{\epsilon_i^T Q_i \epsilon_i + u_{op_i}^T R_{ii} u_{op_i} + \sum_{j=1}^N \alpha_{ij} u_{op_j}^T R_{ij} u_{op_j} - \gamma_{\xi\nu_i}^2 [\omega_i^T S_{1ii}\omega_i \\ &\quad + \sum_{j=1}^N \alpha_{ij}\omega_j^T S_{1ij}\omega_j] - \gamma_{\beta_i}^2 [\beta_i^T S_{2ii}\beta_i + \sum_{j=1}^N \alpha_{ij}\beta_j^T S_{2ij}\beta_j]\} d\tau. \end{aligned} \tag{28}$$

In (28), we have two players: the control player u_{op_i} which tries to minimize the performance index and the disturbance player $\{\omega_i, \beta_i\}$ which tries to maximize it.

Remark 3
Note that in (27), the β_is are considered as arbitrary $\mathcal{L}_2$ disturbances, in contrast to the β_is defined in (24) in Section 3, specified there as part of the feedback (8). This is because the present section

DOI: 10.1002/rnc

primarily seeks to design u_{op_i} with good disturbance suppression properties. Later in Section 5, the actual β_i signals in (24) are accounted for by the small-gain theorem.

The optimal value of this zero-sum graphical game is defined as

$$\begin{aligned} V_i^*(\epsilon_i(0)) &= \min_{u_{op_i}} \quad \max_{\omega_i,\beta_i} J_i(\epsilon_i(0), u_{op_i}, u^*_{op_{-i}}, \omega_i, \omega^*_{-i}, \beta_i, \beta^*_{-i}) \\ &= \max_{\omega_i,\beta_i} \quad \min_{u_{op_i}} J_i(\epsilon_i(0), u_{op_i}, u^*_{op_{-i}}, \omega_i, \omega^*_{-i}, \beta_i, \beta^*_{-i}) \end{aligned} \tag{29}$$

where, $u^*_{op_{-i}}$ and $\{\omega^*_{-i}, \beta^*_{-i}\}$ are the optimal control and disturbance policies of the players in the neighborhood of player i. For the fixed control and disturbance policies u_{op_i} and ω_i, β_i, the quadratic value function is defined as

$$\begin{aligned} V_i(\epsilon_i(t), u_{op_i}, u_{op_{-i}}, \omega_i, \omega_{-i}, \beta_i, \beta_{-i}) = \int_t^{+\infty} \{\epsilon_i^T Q_i \epsilon_i + u_{op_i}^T R_{ii} u_{op_i} + \sum_{j=1}^N \alpha_{ij} u_{op_j}^T R_{ij} u_{op_j} \\ - \gamma^2_{\xi\nu_i}[\omega_i^T S_{1ii}\omega_i + \sum_{j=1}^N \alpha_{ij}\omega_j^T S_{1ij}\omega_j] - \gamma^2_{\beta_i}[\beta_i^T S_{2ii}\beta_i + \sum_{j=1}^N \alpha_{ij}\beta_j^T S_{2ij}\beta_j]\} d\tau. \end{aligned} \tag{30}$$

We make the following assumption regarding the performance index.

Assumption 5
The performance index (28) is zero-state observable.

Assumption 5 guarantees that no solution can identically stay in zero cost other than the zero solution [33]. It is a necessary assumption to prove stability of disturbance-free system under optimal control [31]. Now, we are ready to define the H_∞ graphical game for linear heterogeneous multi-agent systems.

Problem 2 (H_∞ graphical game for linear heterogeneous multi-agent systems)
Consider a group of $N+1$ heterogeneous LTI systems defined by (2-7). Design u_{op_i} to solve optimization problem (29) with respect to dynamics (26).

Remark 4
The graphical game in this paper is a $2N$-player game in the sense that we have N followers each having a disturbance player (ω_i, β_i) and a control player u_i. It is also possible to call this game, a $3N$-player game if we assume for each follower two disturbance players ω_i and β_i separately, and a control player u_i.

In the sequel we solve the H_∞ graphical game problem.

4.2. *Solution of the H_∞ graphical game*

When the value function (30) is finite, using the Leibniz's formula, a differential equivalent to the value function is given in terms of the Hamiltonian

$$H_i = L_i(\epsilon_i, u_{op_i}, u_{op_{-i}}, \omega_i, \omega_{-i}, \beta_i, \beta_{-i}) + \nabla V_i^T \dot{\epsilon}_i = 0 \tag{31}$$

where $\nabla V_i = \frac{\partial V_i}{\partial \epsilon_i}$. At the equilibrium, one has the stationarity conditions

$$\frac{\partial H_i}{\partial u_{op_i}} = \mathbf{0}, \rightarrow u^*_{op_i} = -\frac{1}{2} R_{ii}^{-1} \bar{B}_i^T \nabla V_i, \tag{32}$$

$$\frac{\partial H_i}{\partial \omega_i} = \mathbf{0}, \rightarrow \omega^*_i = \frac{1}{2} \gamma^{-2}_{\xi\nu_i} S_{1ii}^{-1} \bar{P}_i^T \nabla V_i, \tag{33}$$

$$\frac{\partial H_i}{\partial \beta_i} = \mathbf{0}, \rightarrow \beta^*_i = \frac{1}{2} \gamma^{-2}_{\beta_i} S_{2ii}^{-1} \bar{B}_i^T \nabla V_i. \tag{34}$$

Substituting the optimal policy (32) and the worst-case disturbances (33)-(34) into (31) yields coupled Hamilton-Jacobi-Bellman (HJB) equations

$$\begin{aligned}
&\epsilon_i^T Q_i \epsilon_i + \frac{1}{4}\nabla V_i^T \bar{B}_i R_{ii}^{-1} \bar{B}_i^T \nabla V_i + \frac{1}{4}\sum_{j=1}^{N} \alpha_{ij} \nabla V_j^T \bar{B}_j R_{jj}^{-1} R_{ij} R_{jj}^{-1} \bar{B}_j^T \nabla V_j \\
&- \gamma_{\xi\nu_i}^2 \frac{1}{4}[\gamma_{\xi\nu_i}^{-4} \nabla V_i^T \bar{P}_i S_{1ii}^{-1} \bar{P}_i^T \nabla V_i + \sum_{j=1}^{N} \alpha_{ij} \gamma_{\xi\nu_j}^{-4} \nabla V_j^T \bar{P}_j S_{1jj}^{-1} S_{1ij} S_{1jj}^{-1} \bar{P}_j^T \nabla V_j] \\
&- \gamma_{\beta_i}^2 \frac{1}{4}[\gamma_{\beta_i}^{-4} \nabla V_i^T \bar{B}_i S_{2ii}^{-1} \bar{B}_i^T \nabla V_i + \sum_{j=1}^{N} \alpha_{ij} \gamma_{\beta_j}^{-4} \nabla V_j^T \bar{B}_j S_{2jj}^{-1} S_{2ij} S_{2jj}^{-1} \bar{B}_j^T \nabla V_j] \\
&+ \nabla V_i^T [\bar{A}_i \epsilon_i - \frac{1}{2}\bar{B}_i R_{ii}^{-1} \bar{B}_i^T \nabla V_i + \frac{1}{2}\gamma_{\xi\nu_i}^{-2} \bar{P}_i S_{1ii}^{-1} \bar{P}_i^T \nabla V_i + \frac{1}{2}\gamma_{\beta_i}^{-2} \bar{B}_i S_{2ii}^{-1} \bar{B}_i^T \nabla V_i] \\
&- \nabla V_i^T [\sum_{j=1}^{N} \alpha_{ij} \{-\frac{1}{2}\underline{B}_j R_{jj}^{-1} \bar{B}_j^T \nabla V_j + \frac{1}{2}\gamma_{\xi\nu_j}^{-2} \underline{P}_j S_{1jj}^{-1} \bar{P}_j^T \nabla V_j + \frac{1}{2}\gamma_{\beta_j}^{-2} \underline{B}_j S_{2jj}^{-1} \bar{B}_j^T \nabla V_j\}] = 0.
\end{aligned} \tag{35}$$

Based on (35), the HJB equation of player i depends on the HJB equations of other players in its neighborhood. It is in general impossible to solve the coupled HJB equations (35) analytically [31]. Later in Section 6, we use RL techniques and develop a numerical procedure to bring the solutions in real time. For further developments in this section however, we assume the solutions are available.

Let V_i^* be the quadratic optimal solution to (35) and $u_{op_i}^*(V_i^*)$, $\omega_i^*(V_i^*)$ and $\beta_i^*(V_i^*)$ in (32)-(33) be the optimal policy and the worst-case disturbances in terms of V_i^*. In the next theorem, we prove that such V_i^* satisfies the $\mathcal{L}_2$-condition (27).

Theorem 1
Suppose V_i^* is a quadratic positive semi-definite solution to (35) for $i = 1, ..., N$. Let Assumption 5 hold. Using the optimal policy $u_{op_i}^*(V_i^*)$ in (32),

1. The disturbance-free system (26) ($\omega_i \equiv \mathbf{0},\ \beta_i \equiv \mathbf{0},\ i = 1, ..., N$) is asymptotically stable.
2. For all disturbances $\omega_i,\ \omega_{-i},\ \beta_i, \beta_{-i} \in \mathcal{L}_2$, the $\mathcal{L}_2$-condition (27) is satisfied.

Proof
The proof contains two parts. In the first part, we prove the stability of the disturbance-free system (26) and in the second part, we prove the $\mathcal{L}_2$-condition (27).

1. For any smooth value function V_i, the Hamiltonian is defined as

$$\begin{aligned}
&H_i(\epsilon_i, \nabla V_i, u_{op_i}, u_{op_{-i}}, \omega_i, \omega_{-i}, \beta_i, \beta_{-i}) = \\
&\epsilon_i^T Q_i \epsilon_i + u_{op_i}^T R_{ii} u_{op_i} + \sum_{j=1}^{N} \alpha_{ij} u_{op_j}^T R_{ij} u_{op_j} - \gamma_{\xi\nu_i}^2 [\omega_i^T S_{1ii} \omega_i \\
&+ \sum_{j=1}^{N} \alpha_{ij} \omega_j^T S_{1ij} \omega_j] - \gamma_{\beta_i}^2 [\beta_i^T S_{2ii} \beta_i + \sum_{j=1}^{N} \alpha_{ij} \beta_j^T S_{2ij} \beta_j] + \frac{dV_i}{dt}.
\end{aligned} \tag{36}$$

Let V_i^* be a smooth positive semi-definite solution to (35). Completing the squares leads to

$$\begin{aligned}
&H_i(\epsilon_i, \nabla V_i^*, u_{op_i}, u_{op_{-i}}^*, \omega_i, \omega_{-i}^*, \beta_i, \beta_{-i}^*) = \\
&H_i(\epsilon_i, \nabla V_i^*, u_{op_i}^*, u_{op_{-i}}^*, \omega_i^*, \omega_{-i}^*, \beta_i^*, \beta_{-i}^*) + (u_{op_i} - u_{op_i}^*)^T R_{ii} (u_{op_i} - u_{op_i}^*) \\
&\quad - \gamma_{\xi\nu_i}^2 (\omega_i - \omega_i^*)^T S_{1ii} (\omega_i - \omega_i^*) - \gamma_{\beta_i}^2 (\beta_i - \beta_i^*)^T S_{2ii} (\beta_i - \beta_i^*).
\end{aligned}$$

Selecting $u_{op_i} = u^*_{op_i}(V^*_i)$, one has

$$\begin{aligned}&\epsilon_i^T Q_i \epsilon_i + u^{*T}_{op_i} R_{ii} u^*_{op_i} + \sum_{j=1}^{N} \alpha_{ij} u^{*T}_{op_j} R_{ij} u^*_{op_j} - \gamma^2_{\xi\nu_i}[\omega_i^T S_{1ii}\omega_i \\ &+ \sum_{j=1}^{N} \alpha_{ij}\omega_j^T S_{1ij}\omega_j] - \gamma^2_{\beta_i}[\beta_i^T S_{2ii}\beta_i + \sum_{j=1}^{N}\alpha_{ij}\beta_j^T S_{2ij}\beta_j] + \frac{dV^*_i}{dt} \leq 0.\end{aligned} \tag{37}$$

Set $\omega_i \equiv \mathbf{0},\ \omega_{-i} \equiv \mathbf{0},\ \beta_i = \mathbf{0},\ \beta_{-i} \equiv \mathbf{0}$. According to (37)

$$\frac{dV^*_i}{dt} \leq -(\epsilon_i^T Q_i \epsilon_i + u^{*T}_{op_i} R_{ii} u^*_{op_i} + \sum_{j=1}^{N} \alpha_{ij} u^{*T}_{op_j} R_{ij} u^*_{op_j}) \leq 0.$$

Because of Assumption 5, one can conclude from the above inequality that the disturbance-free system (26) is asymptotically stable.

2. Integrate (37)

$$\begin{aligned}&V^*_i(\epsilon_i(T)) - V^*_i(\epsilon_i(0)) + \int_0^T \{\epsilon_i^T Q_i \epsilon_i + u^{*T}_{op_i} R_{ii} u^*_{op_i} + \sum_{j=1}^{N} \alpha_{ij} u^{*T}_{op_j} R_{ij} u^*_{op_j} - \gamma^2_{\xi\nu_i}[\omega_i^T S_{1ii}\omega_i \\ &+ \sum_{j=1}^{N} \alpha_{ij}\omega_j^T S_{1ij}\omega_j] - \gamma^2_{\beta_i}[\beta_i^T S_{2ii}\beta_i + \sum_{j=1}^{N}\alpha_{ij}\beta_j^T S_{2ij}\beta_j]\} d\tau \leq 0.\end{aligned}$$

Select $\epsilon_i(0) = \mathbf{0}$. Since $V^*_i(\epsilon_i(0)) = 0$ and $V^*_i(\epsilon_i(T)) > 0$, (27) is satisfied.

□

Remark 5
Assuming a quadratic structure for the value function $V^*_i = 0.5\epsilon_i^T P^g_i \epsilon_i$, the control signal (32) reads

$$u_{op_i} = -R_{ii}^{-1}\bar{B}_i^T P_i^g \epsilon_i = -R_{ii}^{-1}\bar{B}_i^T P_i^g \begin{bmatrix} e_{\xi_i} \\ \nu_i \end{bmatrix} = K_{i1} e_{\xi_i} + K_{i2}\nu_i.$$

The above equation is useful in finding the controller gain K_i. Recall the definition of policy u_{op_i} in (25)

$$u_{op_i} = K_i R_2 e_{\xi_i} + (d_i + g_i)\bar{K}_i \nu_i.$$

In order to calculate K_i from $K_{i1} = K_i R_2$, the following standard condition is required

$$Rank(R_2) = Rank\left(\begin{bmatrix} R_2 \\ K_{i1}\end{bmatrix}\right). \tag{38}$$

One can satisfy (38) by e.g. selecting any invertible R_2 similar to [2, 3]. Then, $K_i = K_{i1}R_2^{-1}$ and $\bar{K}_i = \frac{1}{(d_i+g_i)}K_{i2}$. For convenience we assume,

Assumption 6
R_2 is invertible.

4.3. Nash equilibrium solution

In an H_∞ graphical game, we are interested in convergence to a Nash equilibrium (NE). The game has a well-defined NE if

$$\begin{aligned} J_i(u^*_{op_i}, u^*_{op_{-i}}, \omega_i, \omega^*_{-i}, \beta_i, \beta^*_{-i}) &\leq J_i(u^*_{op_i}, u^*_{op_{-i}}, \omega^*_i, \omega^*_{-i}, \beta^*_i, \beta^*_{-i}) \\ &\leq J_i(u_{op_i}, u^*_{op_{-i}}, \omega^*_i, \omega^*_{-i}, \beta^*_i, \beta^*_{-i}). \end{aligned} \tag{39}$$

The following theorem shows that a positive solution to (35) also guarantees convergence to an NE.

DOI: 10.1002/rnc

Theorem 2
Let V_i^* be a quadratic positive semi-definite solution to (35) for $i = 1, ..., N$ such that the closed-loop system of

$$\dot{\epsilon}_i = \bar{A}_i \epsilon_i + \bar{B}_i u^*_{op_i} + \bar{P}_i \omega^*_i + \bar{B}_i \beta^*_i - \sum_{j=1}^{N} \alpha_{ij} (\underline{B}_j u^*_{op_j} + \underline{P}_j \omega^*_{op_j} + \underline{B}_j \beta^*_j) \tag{40}$$

is asymptotically stable. Let Assumption 5 hold. Then,

1. The control signal and disturbances (32)-(34) form an NE.
2. The optimal value of the game is then $V_i^*(\epsilon_i(0))$.
3. The optimal control (32) solves Problem 2.

Proof

1. First, we show that (32)-(34) form an NE. Rewriting (28) and adding zero

$$\begin{aligned} J_i &= \int_0^{+\infty} L_i d\tau + \int_0^{+\infty} \dot{V}_i d\tau + V_i(\epsilon_i(0)) - V_i(\epsilon_i(+\infty)) \\ &= \int_0^{+\infty} H_i(\epsilon_i, \nabla V_i, u_{op_i}, u_{op_{-i}}, \omega_i, \omega_{-i}, \beta_i, \beta_{-i}) d\tau + V_i(\epsilon_i(0)) - V_i(\epsilon_i(+\infty)). \end{aligned}$$

Let V_i^* be a smooth positive semi-definite solution to (35). By completing the squares one has

$$\begin{aligned} J_i =& V_i(\epsilon_i(0)) - V_i(\epsilon_i(+\infty)) + \int_0^{+\infty} \{ H_i(\epsilon_i, \nabla V_i^*, u^*_{op_i}, u^*_{op_{-i}}, \omega^*_i, \omega^*_{-i}, \beta^*_i, \beta^*_{-i}) \\ &+ (u_{op_i} - u^*_{op_i})^T R_{ii} (u_{op_i} - u^*_{op_i}) + \sum_{j=1}^{N} \alpha_{ij} (u_{op_j} - u^*_{op_j})^T R_{ij} (u_{op_j} - u^*_{op_j}) \\ &+ 2 \sum_{j=1}^{N} \alpha_{ij} u^{*T}_{op_j} R_{ij} (u_{op_j} - u^*_{op_j}) - \sum_{j=1}^{N} \alpha_{ij} \nabla V_i^T \underline{B}_j (u_{op_j} - u^*_{op_j}) \\ &- \gamma^2_{\xi\nu_i} (\omega_i - \omega^*_i)^T S_{1ii} (\omega_i - \omega^*_i) - \gamma^2_{\xi\nu_i} \sum_{j=1}^{N} \alpha_{ij} (\omega_j - \omega^*_j)^T S_{1ij} (\omega_j - \omega^*_j) \\ &- 2\gamma^2_{\xi\nu_i} \sum_{j=1}^{N} \alpha_{ij} \omega^{*T}_j S_{1ij} (\omega_j - \omega^*_j) - \sum_{j=1}^{N} \alpha_{ij} \nabla V_i^T \underline{P}_j (\omega_j - \omega^*_j) \\ &- \gamma^2_{\beta_i} (\beta_i - \beta^*_i)^T S_{2ii} (\beta_i - \beta^*_i) - \gamma^2_{\beta_i} \sum_{j=1}^{N} \alpha_{ij} (\beta_j - \beta^*_j)^T S_{2ij} (\beta_j - \beta^*_j) \\ &- \gamma^2_{\beta_i} \sum_{j=1}^{N} \alpha_{ij} \beta^{*T}_j S_{2ij} (\beta_j - \beta^*_j) - \sum_{j=1}^{N} \alpha_{ij} \nabla V_i^T \underline{B}_j (\beta_j - \beta^*_j) \} d\tau. \end{aligned}$$

Since (40) is asymptotically stable, $V_i^*(\epsilon_i(+\infty)) = 0$. Set $u_{op_j} = u^*_{op_j}$, $\omega_j = \omega^*_j$, $\beta_j = \beta^*_j$, $\forall j \in N_i$ in the above equation. Then (39) follows easily.
2. Next, we obtain the optimal value of the game. Set $u_{op_i} = u^*_{op_i}$, $\omega_i = \omega^*_i$, $\beta_i = \beta^*_i$; then, $J_i^* = V_i^*(\epsilon_i(0))$.
3. It is concluded from part 2.

□

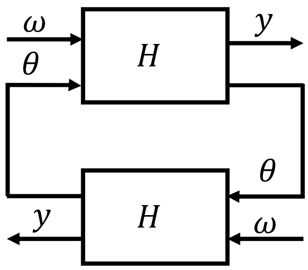


Figure 1. The interconnected system

Remark 6
Theorem 2 shows that the controls $u^*_{op_i}$ in (32) lead to the desired NE satisfying the $\mathcal{L}_2$-condition (27) and thus solve Problem 2. Note however that Problem 2, and indeed all the developments of this section, consider u_{op_i} and β_i as independent signals, whereas those are in fact inseparably linked in the actual control signal (20). This fact requires additional conditions, elaborated further in Section 5, guaranteeing that the control design proposed here solves the original Problem 1 of Section 3.

5. H_∞ OUTPUT REGULATION USING GRAPHICAL GAME

In Section 4, we shown that V_i^* in (35) solves Problem 2, (see Theorem 2). Building on developments of Section 4, in this section, we give conditions guaranteeing that the graphical game solution V_i^* and the control designed from it can indeed be used to solve the H_∞ output regulation problem (Problem 1). Moreover, we give an upper bound for the overall $\mathcal{L}_2$-gain of output synchronization error with respect to disturbances; *i.e.* $L_{\delta_y \omega}$.

Theorem 3
Let H_1 and H_2 be two subsystems, whose inputs are θ_1, ω and θ_2, ω and outputs are y_1, y_2, respectively. Let $\omega_1,\ \omega_2 \in \mathcal{L}_2$, and assume that H_1 and H_2 are interconnected as depicted in Fig. 1 and

$$\|y_1\|_{\mathcal{L}_2} \leq L_{11}\|\omega\|_{\mathcal{L}_2} + L_{12}\|\theta_1\|_{\mathcal{L}_2}, \quad \|y_2\|_{\mathcal{L}_2} \leq L_{22}\|\omega\|_{\mathcal{L}_2} + L_{21}\|\theta_2\|_{\mathcal{L}_2}. \tag{41}$$

If $L_{12}L_{21} < 1$, then $y_1 \in \mathcal{L}_2$ and

$$\|y_1\|_{\mathcal{L}_2} \leq (1 - L_{12}L_{21})^{-1}(L_{11} + L_{12}L_{22})\|\omega\|_{\mathcal{L}_2}. \tag{42}$$

Proof
According to (41)

$$\begin{aligned}\|y_1\|_{\mathcal{L}_2} &\leq L_{11}\|\omega\|_{\mathcal{L}_2} + L_{12}\|\theta_1\|_{\mathcal{L}_2} \leq L_{11}\|\omega\|_{\mathcal{L}_2} + L_{12}\|y_2\|_{\mathcal{L}_2} \\ &\leq L_{11}\|\omega\|_{\mathcal{L}_2} + L_{12}(L_{22}\|\omega\|_{\mathcal{L}_2} + L_{21}\|y_1\|_{\mathcal{L}_2}).\end{aligned}$$

Hence

$$(1 - L_{12}L_{21})\|y_1\|_{\mathcal{L}_2} \leq (L_{11} + L_{12}L_{22})\|\omega\|_{\mathcal{L}_2}.$$

By the small gain theorem, the $\mathcal{L}_2$-gain stability is guaranteed if $L_{12}L_{21} < 1$ and (42) follows. □

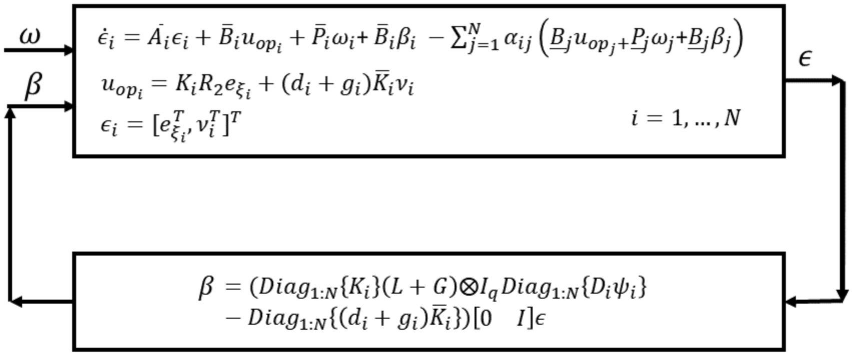


Figure 2. The representation of the nominal and the interconnected systems

The overall system of (24)-(26) can be represented as in Fig. 2. There, we have two interconnected subsystems. We name the upper block subsystem *the nominal system*, whose description is given in (25)-(26) and we call the lower block subsystem *the interconnected system* whose description is given in (24). This representation is useful as we can use Theorem 3 to prove $\mathcal{L}_2$-stability. The following theorem presents one of the main results of this paper; it specifies the set of conditions such that V_i^* in (35) and the control based on it solve the H_∞ output regulation problem (Problem 1). Define

$$
\begin{aligned}
S_{m_i} &:= \max_{i,j\in N_i}\{\bar{\sigma}(\sqrt{\alpha_{ij}S_{1ij}}), \bar{\sigma}(\sqrt{S_{1ii}})\}, \quad \beta_{m_i} := \max_{i,j\in N_i}\{\bar{\sigma}(\sqrt{\alpha_{ij}S_{2ij}}), \bar{\sigma}(\sqrt{S_{2ii}})\},\\
L_y &= \frac{1}{\underline{\sigma}(T_g)}\left[\max_i \bar{\sigma}(C_iT_i) + \max_i \bar{\sigma}(E_iK_i)\bar{\sigma}(L+G)\max_i \bar{\sigma}(D_iT_i)\right],\\
L_{21} &= \max_i \bar{\sigma}(K_i)\bar{\sigma}(L+G)\max_i \bar{\sigma}(D_i\Psi_i) + \max_i \bar{\sigma}((d_i+g_i)\bar{K}_i)\\
L_{11} &= \sqrt{\sum_{i=1}^{N}\underline{\sigma}(\sqrt{Q_i})^{-2}\gamma_{\xi\nu_i}^2 S_{m_i}^2}, \quad L_{12} = \sqrt{\sum_{i=1}^{N}\underline{\sigma}(\sqrt{Q_i})^{-2}\gamma_{\beta_i}^2 \beta_{m_i}^2},\\
T_g &= \begin{bmatrix}(L+G)\otimes I_l & \mathbf{0}\\ \mathbf{0} & I_{\sum_{i=1}^{N}n_i - Nl}\end{bmatrix}.
\end{aligned}
\tag{43}
$$

Theorem 4
Let Assumptions 1-5 hold. Let V_i^* be a quadratic positive semi-definite solution to (35) and $u_{op_i}^*(V_i^*)$ in (32) be the optimal policy. Assume that the controller gain K_i is obtained according to Remark 5. If either (i) $\bar{\sigma}(K_i)\bar{\sigma}(D_i\Psi_i)$ and $\bar{\sigma}(\bar{K}_i)$ are sufficiently small or (ii) γ_{β_i} is sufficiently small for $i = 1, ..., N$, then,

1. An upper bound for the $\mathcal{L}_2$-gain of output synchronization error δ_y with respect to ω is given by
$$L_{\delta_y\omega} = L_y(1 - L_{12}L_{21})^{-1}L_{11}. \tag{44}$$
2. Distributed control (8) solves the H_∞ output regulation problem (Problem 1).

Proof

DOI: 10.1002/rnc

1. Consider subsystem H_1 from Theorem 3 as the overall system of (26). This subsystem has two inputs ω and β, where ω is the disturbance input and β is related to the subsystem H_2, and one output ϵ. First note that we always have

$$\int_0^T \underline{\sigma}(\sqrt{Q_i})^2\|\epsilon_i\|_2^2 d\tau \leq \int_0^T \epsilon_i^T Q_i \epsilon_i + u_{op_i}^T R_{ii} u_{op_i} + \sum_{j=1}^N \alpha_{ij} u_{op_j}^T R_{ij} u_{op_j} d\tau.$$

Using the above inequality in accordance with (27),

$$\int_0^T \|\epsilon\|_2^2 d\tau = \int_0^T \sum_{i=1}^N \|\epsilon_i\|_2^2 d\tau$$
$$\leq \sum_{i=1}^N \underline{\sigma}(\sqrt{Q_i})^{-2}\gamma_{\xi\nu_i}^2 S_{m_i}^2 \int_0^T \|\omega\|_2^2 d\tau + \sum_{i=1}^N \underline{\sigma}(\sqrt{Q_i})^{-2}\gamma_{\beta_i}^2 \beta_{m_i}^2 \int_0^T \|\beta\|_2^2 d\tau.$$

Hence, the $\mathcal{L}_2$-gain of ϵ with respect to ω and β reads

$$\|\epsilon\|_{\mathcal{L}_2} \leq \sqrt{\sum_{i=1}^N \underline{\sigma}(\sqrt{Q_i})^{-2}\gamma_{\xi\nu_i}^2 S_{m_i}^2}\,\|\omega\|_{\mathcal{L}_2} + \sqrt{\sum_{i=1}^N \underline{\sigma}(\sqrt{Q_i})^{-2}\gamma_{\beta_i}^2 \beta_{m_i}^2}\,\|\beta\|_{\mathcal{L}_2}. \quad (45)$$

Next, we consider the subsystem H_2 from Theorem 3 as (24) with input ϵ and output β. According to the definition of β in (24), the $\mathcal{L}_2$-gain of β with respect to ϵ reads

$$\|\beta\|_{\mathcal{L}_2} \leq (\max_i \bar{\sigma}((d_i+g_i)\bar{K}_i) + \max_i \bar{\sigma}(K_i)\bar{\sigma}(L+G)\max_i \bar{\sigma}(D_i\Psi_i))\|\epsilon\|_{\mathcal{L}_2}. \quad (46)$$

By substituting (46) in (45)

$$\|\epsilon\|_{\mathcal{L}_2} \leq (1 - L_{12}L_{21})^{-1} L_{11}\|\omega\|_{\mathcal{L}_2}. \quad (47)$$

By Theorem 3, $\epsilon \in \mathcal{L}_2$ if $L_{12}L_{21} < 1$ which is satisfied by either of the conditions (i)-(ii) in the body of the theorem. Let $\bar{x}_i = [\delta_{\xi_i}^T, \nu_i^T]$ where $\delta_{\xi_i} = \xi_i - \xi_0$. Let ϵ and $\bar{x}$ be the overall vectors of ϵ_i and $\bar{x}_i$ respectively. Then,

$$\epsilon = T_g \bar{x}, \quad (48)$$

where T_g is given in (43). The output synchronization error δ_{y_i} in the coordinate $\bar{x}_i$ reads

$$\delta_{y_i} = C_i x_i + E_i u_i - R_1\xi_0 = C_i T_i \bar{x}_i + E_i K_i \left[\sum_{i=1}^N \alpha_{ij}(D_i T_i \bar{x}_i - D_j T_j \bar{x}_j) + g_i D_i T_i \bar{x}_i\right].$$

Using (47)-(48) and the above equation, the $\mathcal{L}_2$-gain of the output synchronization error δ_y reads

$$\|\delta_y\|_{\mathcal{L}_2} \leq L_y(1 - L_{12}L_{21})^{-1} L_{11}\|\omega\|_{\mathcal{L}_2}, \quad (49)$$

which gives the upper bound in (44).

2. To prove that (8) solves the H_∞ output regulation problem, we need to show that the three properties in Problem 1 hold.

   a. By (47)-(48), $\bar{x} \in \mathcal{L}_2$. Because of the linearity of the system and by setting $\omega \equiv \mathbf{0}$, one can conclude that $\bar{x} \to \mathbf{0}$. Using the definition of $\bar{x}_i^T = [\xi_i - \xi_0, \nu_i]$, by setting $\xi_0 \equiv \mathbf{0}$, we have $x_i \to \mathbf{0}$.

   b. By (49), $\delta_y \in \mathcal{L}_2$. By linearity of the system and $\omega \equiv \mathbf{0}$, we conclude that $\delta_y \to \mathbf{0}$. Similarly, we can show that $\delta_z \to \mathbf{0}$.

c. The $\mathcal{L}_2$-gain of δ_y with respect to ω is given in (49) and its finiteness is guaranteed by the conditions (i) or (ii) in the body of the theorem.

$\square$

Remark 7
In Section 4, the designed control u_{op_i} and the signal β_i were considered separate while they are linked in the actual control signal (20) owing to heterogeneity of the agents. Theorem 4 reconciles development of graphical games in Section 4 with the original problem and takes the linkage of u_{op_i} and β_i in the actual control signal (20) into consideration. This requires additional conditions (given in the body of Theorem 4) that do not appear for homogeneous agents. Note that conclusions of Theorem 4 indeed reduce for identical agents to the cases familiar from the literature [20, 21]. Note also that here the graph is not required to be acyclic [34].

6. ONLINE SOLUTION TO H_∞ GRAPHICAL GAME

As we dicussed in Section 4, one needs to obtain solutions to the coupled partial differential HJB equations (35) to solve the graphical game problem. It is in general impossible to solve these equations analytically. However, RL has shown promising results in solving such complicated coupled equations numerically and is the only viable method applicable in real time. In this section, we propose a numerical RL procedure to design the controller gain K_i and to obtain solutions to the coupled HJB equations (35) in real time. Note that because of the heterogeneity of the agents in our paper, the RL frameworks in [20, 21] cannot be used.

Our online learning structure uses four adaptive networks. The first network approximates the value function and is named the critic network. The second one approximates the control policy u_{op_i} in (32) and is named the actor network. The third and fourth networks approximate the disturbances ω_i and β_i in (33)-(34).

Assume that the value function $V_i(\epsilon_i(t))$ is smooth. Then, according to Weierstrass higher-order approximation theorem, one can approximate $V_i(\epsilon_i(t))$ by

$$V_i(\epsilon_i) = W_i^T \Phi_i(\epsilon_i) + \varepsilon_i,$$

in which, Φ_i is a basis function vector with μ_{n_i} neurons, W_i is the optimal weight and ε_i is the approximation error. The weights of the critic network, which provide the best approximation to (35), are unknown. Let $\hat{W}_i$ denote the current estimate of the critic weights. Then, the approximated value function is given by

$$\hat{V}_i(\epsilon_i) = \hat{W}_i^T \Phi_i(\epsilon_i), \tag{50}$$

and the approximated error of the Bellman equation is

$$H_i = L_i + \hat{W}_i^T \nabla\Phi_i \dot{\epsilon}_i = \varepsilon_{H_i}. \tag{51}$$

Next, we define an actor network to approximate the optimal policy (32)

$$\hat{u}_{op_i} = -\frac{1}{2} R_{ii}^{-1} \bar{B}_i^T \nabla\Phi_i^T \hat{W}_{i+N}, \tag{52}$$

where $\hat{W}_{i+N}$ denotes the current estimate of the actor weights. Also, we use two additional networks to approximate the worst-case disturbances (33)-(34)

$$\hat{\omega}_i = \frac{1}{2}\gamma_{\xi\nu_i}^{-2} S_{1ii}^{-1} \bar{P}_i^T \nabla\Phi_i^T \hat{W}_{i+2N}, \tag{53}$$

$$\hat{\beta}_i = \frac{1}{2}\gamma_{\beta_i}^{-2} S_{2ii}^{-1} \bar{B}_i^T \nabla\Phi_i^T \hat{W}_{i+3N}, \tag{54}$$

where $\hat{W}_{i+2N}$ and $\hat{W}_{i+3N}$ denote the current estimates of the weights of disturbances ω_i and β_i respectively.

The following theorem presents another main result of this paper; it gives the tuning laws for the adaptive network weights such that the HJB equation (35) is solved numerically, and closed-loop system of (26), the weight estimation errors $\tilde{W}_i = W_i - \hat{W}_i$, $\tilde{W}_{i+N} = W_i - \hat{W}_{i+N}$, $\tilde{W}_{i+2N} = W_i - \hat{W}_{i+2N}$, $\tilde{W}_{i+3N} = W_i - \hat{W}_{i+3N}$ are locally Uniformly Ultimately Bounded (UUB) [20].

Theorem 5

Consider Problem 2 and let the conditions in Theorem 4 hold. Assume that the value function, the policy and the disturbances are estimated by (50) and (52)-(54) respectively. Assume that

$$\sigma_{i+N} = \nabla\Phi_i\{\bar{A}_i\epsilon_i + \bar{B}_i\hat{u}_{op_i} + \bar{P}_i\hat{\omega}_i + \bar{B}_i\hat{\beta}_i - \sum_{j=1}^{N}\alpha_{ij}(\underline{B}_j\hat{u}_{op_j} + \underline{P}_j\hat{\omega}_j + \underline{B}_j\hat{\beta}_j)\}, \quad (55)$$

is Persistently Exciting (PE). Tune the weights of the critic network as

$$\dot{\hat{W}}_i = -a_i\frac{\sigma_{i+N}}{(1+\sigma_{i+N}^T\sigma_{i+N})^2}\varepsilon_{H_i}^{op} \quad (56)$$

where $\varepsilon_{H_i}^{op}$ is obtained by inserting (52)-(54) in (51). Tune the weights of the actor and disturbance networks as

$$\begin{aligned}\dot{\hat{W}}_{i+N} =& - a_{i+N}\{(\bar{S}_{1i}\hat{W}_{i+N} - \bar{T}_{1i}\hat{W}_i) - \frac{1}{4}\bar{D}_i\hat{W}_{i+N}\frac{\bar{\sigma}_{i+N}^T}{m_{s_i}}\hat{W}_i \\ & - \frac{1}{4}\sum_{j=1}^{N}\alpha_{ji}\nabla\Phi_i\bar{B}_iR_{ii}^{-1}R_{ji}R_{ii}^{-1}\bar{B}_i^T\nabla\Phi_i^T\hat{W}_{i+N}\frac{\bar{\sigma}_{j+N}^T}{m_{s_j}}\hat{W}_j\},\end{aligned} \quad (57)$$

$$\begin{aligned}\dot{\hat{W}}_{i+2N} =& - a_{i+2N}\{(\bar{S}_{2i}\hat{W}_{i+2N} - \bar{T}_{2i}\hat{W}_i) + \frac{1}{4}\gamma_{\xi\nu_i}^{-2}\bar{E}_{2i}\hat{W}_{i+2N}\frac{\bar{\sigma}_{i+N}^T}{m_{s_i}}\hat{W}_i \\ & + \frac{1}{4}\gamma_{\xi\nu_i}^{-2}\sum_{j=1}^{N}\alpha_{ji}\nabla\Phi_i\bar{P}_iS_{1ii}^{-1}S_{1ji}S_{1ii}^{-1}\bar{P}_i^T\nabla\Phi_i^T\hat{W}_{i+2N}\frac{\bar{\sigma}_{j+N}^T}{m_{s_j}}\hat{W}_j\},\end{aligned} \quad (58)$$

$$\begin{aligned}\dot{\hat{W}}_{i+3N} =& - a_{i+3N}\{(\bar{S}_{3i}\hat{W}_{i+3N} - \bar{T}_{3i}\hat{W}_i) + \frac{1}{4}\gamma_{\beta_i}^{-2}\bar{E}_{3i}\hat{W}_{i+3N}\frac{\bar{\sigma}_{i+N}^T}{m_{s_i}}\hat{W}_i \\ & + \frac{1}{4}\gamma_{\beta_i}^{-2}\sum_{j=1}^{N}\alpha_{ji}\nabla\Phi_i\bar{B}_iS_{2ii}^{-1}S_{2ji}S_{2ii}^{-1}\bar{B}_i^T\nabla\Phi_i^T\hat{W}_{i+3N}\frac{\bar{\sigma}_{j+N}^T}{m_{s_j}}\hat{W}_j\},\end{aligned} \quad (59)$$

where,

$$\begin{aligned}&\bar{\sigma}_{i+N} = \frac{\sigma_{i+N}}{1+\sigma_{i+N}^T\sigma_{i+N}}, \quad m_{s_i} = 1+\sigma_{i+N}^T\sigma_{i+N}, \\ &\bar{D}_i = \nabla\Phi_i\bar{B}_iR_{ii}^{-1}\bar{B}_i^T\nabla\Phi_i^T, \quad \bar{E}_{2i} = \nabla\Phi_i\bar{P}_iS_{1ii}^{-1}\bar{P}_i^T\nabla\Phi_i^T, \quad \bar{E}_{3i} = \nabla\Phi_i\bar{B}_iS_{2ii}^{-1}\bar{B}_i^T\nabla\Phi_i^T,\end{aligned} \quad (60)$$

and $a_i > 0,\ a_{i+N} > 0,\ a_{i+2N} > 0,\ \bar{T}_{1i} > 0,\ \bar{S}_{1i} > 0,\ \bar{T}_{2i} > 0,\ \bar{S}_{2i} > 0,\ \bar{T}_{3i} > 0,\ \bar{S}_{3i} > 0,\ i = 1, ..., N$ are the tuning parameters. Then,

1. The closed-loop system of (26), the weight estimation errors $\tilde{W}_i$, $\tilde{W}_{i+N}$, $\tilde{W}_{i+2N}$ and $\tilde{W}_{i+3N}$ are locally UUB.
2. ε_{H_i} is UUB and $\hat{W}_i$ converges to the approximated solution of the HJB equation (35) for $i = 1, ..., N$.
3. $\hat{u}_{op_i}$, $\hat{\omega}_i$, $\hat{\beta}_i$ converge to the approximated NE.

Proof

The proof is similar as in [19]. □

DOI: 10.1002/rnc

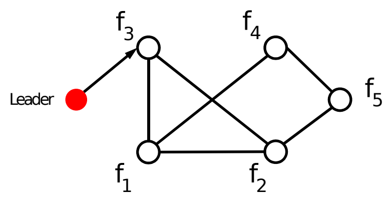


Figure 3. The communication graph

Remark 8

The tuning laws (56)-(59) in Theorem 5 are fully distributed and they depend only on local information available to each single-agent. Hence those are indeed applicable on undirected graphs, satisfying Assumption 1. The tuning laws (56)-(57) bring solutions to coupled HJB equations (35) and the applicable controls (20) in real time for heterogeneous agents. Additional adaptive networks (54) are used to estimate the β_i signals stemming from agent heterogeneity. These networks are absent for homogeneous agents, see [20, 21].

7. SIMULATION RESULTS

Consider a group of five followers and one leader communicating with each other according to the graph shown in Fig. 3. The edge weights in the communication graph are all set to one. Consider the followers' dynamics as

$$\begin{cases} \dot{x}_i = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & -7 & 2 \\ 0 & 0 & 0 & -4 \end{bmatrix} x_i + \begin{bmatrix} 0 \\ 1 \\ -0.2 \\ 0 \end{bmatrix} u_i + \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \omega_i \\ y_i = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} x_i, \quad z_i = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0.005 & 0 \end{bmatrix} x_i, \quad \text{for} \quad i = 1,3,5, \end{cases}$$

$$\begin{cases} \dot{x}_i = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -2 \end{bmatrix} x_i + \begin{bmatrix} 0.1 \\ 1 \\ 0.2 \end{bmatrix} u_i + \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \omega_i, \\ y_i = \begin{bmatrix} 1 & 1 & -1 \end{bmatrix} x_i, \quad z_i = \begin{bmatrix} 1 & 0 & 0.001 \\ 0 & 1 & -0.003 \end{bmatrix} x_i, \quad \text{for} \quad i = 2,4. \end{cases}$$

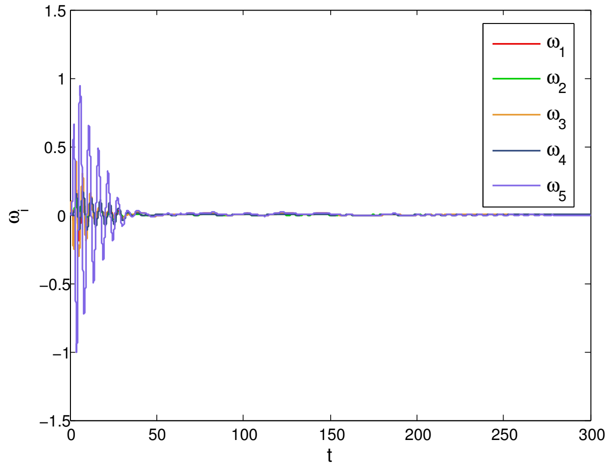
Consider the leader's dynamics as

$$\dot{\xi}_0 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \xi_0, \quad y_0 = \begin{bmatrix} 1 & 1 \end{bmatrix} \xi_0, \quad z_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \xi_0.$$

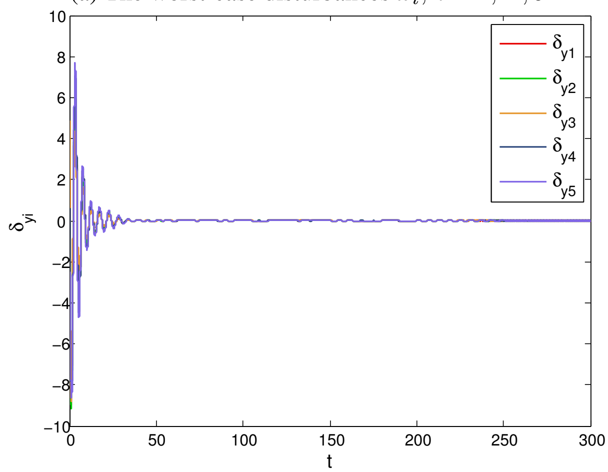
According to the followers' and the leader's dynamics one can find that

$$\Pi_i = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \ \Psi_i = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{for } i = 1,3,5, \text{ and } \quad \Pi_i = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \ \Psi_i = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \text{for } i = 2,4.$$

DOI: 10.1002/rnc



(a) The worst-case disturbances $\omega_i,\ i = 1, ..., 5$



(b) y-output synchronization error $\delta_{yi},\ i = 1, ..., 5$

Figure 4. H_∞ output regulation results

We use four adaptive networks as detailed in Section 6 to solve the differential graphical game for H_∞ output regulation in real time. The weights in the $\mathcal{L}_2$-condition (27) are selected as

$$Q_i = 10I, \quad S_{1ii} = I, \quad S_{2ii} = I, \quad R_{ii} = 0.5, \quad S_{1ij} = 0.05I, \quad S_{2ij} = 0.05I$$
$$R_{ij} = 0.05, \quad \gamma_{\xi\nu_i} = 1.5, \quad \gamma_{\beta_i} = 0.8, \quad i = 1, ..., 5, \quad j \in N_i.$$

We select the tuning parameters as

$$a_i = 0.5,\ a_{i+N} = a_{i+2N} = a_{i+3N} = 0.05,\ \bar{S}_{1i} = \bar{S}_{2i} = \bar{S}_{3i} = I,\ \bar{T}_{1i} = \bar{T}_{2i} = \bar{T}_{3i} = I,\ i = 1, ..., 5.$$

The graphical game is implemented according to Theorem 5 and the gain K_i is obtained as shown in Remark 5. The worst-case disturbances in (53) are applied to the agents and are shown in Fig. 4a. The y-output synchronization errors for followers 1-5 are shown in Fig. 4b. Our simulation example illustrates feasibility and efficiency of adaptive networks for solving coupled HJB equations (35)-those networks perform sufficiently fast to be run online for reasonably large systems.

8. CONCLUSION

In this paper, we have defined H_∞ graphical games for linear heterogeneous agents. Because of heterogeneity of the agents, we have used the output regulation theory to define the node error

DOI: 10.1002/rnc

dynamics properly. Our graphical games formulation has two important properties: firstly, the agents are heterogeneous and secondly, unmodeled disturbances are present. This allows us to solve output regulation and H_∞ output regulation problems. We have obtained an upper bound for the $\mathcal{L}_2$-gain of the output synchronization error with respect to disturbances which has not been hitherto derived in the graphical game framework. We believe that the current developments shed some light on the problem of graphical games with heterogeneous agents.

ACKNOWLEDGEMENT

Farnaz Adib Yaghmaie is supported by the Vinnova Competence Center LINK-SIC and by the Wallenberg Artificial Intelligence, Autonomous Systems and Software Program (WASP). Kristian Hengster Movric is supported by the GACR grant 16-25493Y. Frank l. Lewis is supported by ONR grant N00014-17-1-2239, ONR Grant N00014-18-1-2221, NSF Grant ECCS-1839804, and China NSFC grant #61633007. Rong Su is supported by NRF BCA GBIC grant on Scalable and Smart Building Energy Management (NRF2015ENC-GBICRD001-057) and MoE Academic Research Grant on Secure and Privacy Preserving Multi-Agent Cooperation (RG94/17-(S)-SU RONG (VP)).

REFERENCES

1. K. Hengster-Movric, F. L. Lewis, and M. Sebek, “Distributed static output-feedback control for state synchronization in networks of identical LTI systems,” *Automatica*, vol. 53, pp. 282–290, 2015.
2. C.-Q. Ma and J.-F. Zhang, “Necessary and sufficient conditions for consensusability of linear multi-agent systems,” *IEEE Transactions on Automatic Control*, vol. 55, no. 5, pp. 1263–1268, 2010.
3. L. Scardovi and R. Sepulchre, “Synchronization in networks of identical linear systems,” *Automatica*, vol. 45, no. 11, pp. 2557–2562, 2009.
4. F. Adib Yaghmaie, F. L. Lewis, and R. Su, “Output regulation of linear heterogeneous multi-agent systems via output and state feedback,” *Automatica*, vol. 67, pp. 157–164, 2016.
5. P. Wieland, R. Sepulchre, and F. Allgöwer, “An internal model principle is necessary and sufficient for linear output synchronization,” *Automatica*, vol. 47, no. 5, pp. 1068–1074, 2011.
6. C. Huang and X. Ye, “Cooperative output regulation of heterogeneous multi-agent systems: An H_∞ criterion,” *Automatic Control, IEEE Transactions on*, vol. 59, pp. 267–273, 2014.
7. N. Yang, J.-W. Xiao, L. Xiao, and Y.-W. Wang, “Non-zero sum differential graphical game: cluster synchronisation for multi-agents with partially unknown dynamics,” *International Journal of Control*, pp. 1–12, 2018.
8. F. Adib Yaghmaie, R. Su, F. L. Lewis, and L. Xie, “Multi-party consensus of linear heterogeneous multi-agent systems,” *IEEE Transactions on Automatic Control*, vol. 62(11), pp. 5578 – 5589, 2017.
9. Y.-W. Wang, X.-K. Liu, J.-W. Xiao, and Y. Shen, “Output formation-containment of interacted heterogeneous linear systems by distributed hybrid active control,” *Automatica*, vol. 93, pp. 26–32, 2018.
10. X. Wang, Y. Hong, J. Huang, and Z.-P. Jiang, “A distributed control approach to a robust output regulation problem for multi-agent linear systems,” *Automatic Control, IEEE Transactions on*, vol. 55, no. 12, pp. 2891–2895, 2010.
11. J. Lunze, “Synchronization of heterogeneous agents,” *IEEE transactions on automatic control*, vol. 57, no. 11, pp. 2885–2890, 2012.
12. F. Adib Yaghmaie, F. L. Lewis, and R. Su, “Output regulation of heterogeneous linear multi-agent systems with differential graphical game,” *International Journal of Robust and Nonlinear Control*, vol. 26, pp. 2256–2278, 2016.
13. Y. Liu and Y. Jia, “Robust H_∞ consensus control of uncertain multi-agent systems with time delays,” *International Journal of Control, Automation and Systems*, vol. 9, no. 6, pp. 1086–1094, 2011.
14. X. Yang and J. Wang, “Finite-gain $\mathcal{L}_p$ consensus of multi-agent systems,” *International Journal of Control, Automation and Systems*, vol. 11, no. 4, pp. 666–674, 2013.
15. F. Adib Yaghmaie, K. Hengster Movric, F. L. Lewis, R. Su, and M. Sebek, “H_∞-output regulation of linear heterogeneous multiagent systems over switching graphs,” *International Journal of Robust and Nonlinear Control*, vol. 28, no. 13, pp. 3852–3870, 2018.
16. Q. Jiao, H. Modares, F. L. Lewis, S. Xu, and L. Xie, “Distributed l2-gain output-feedback control of homogeneous and heterogeneous systems,” *Automatica*, vol. 71, pp. 361–368, 2016.
17. M. Zhang, A. Saberi, and A. A. Grip, Havard F.and Stoorvogel, “H_∞ almost output synchronization for heterogeneous networks without exchange of controller states,” *IEEE transactions on control of network systems*, vol. 2, no. 4, pp. 348–357, 2015.
18. D. Liu, H. Li, and D. Wang, “Online synchronous approximate optimal learning algorithm for multi-player non-zero-sum games with unknown dynamics,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 44, no. 8, pp. 1015–1027, 2014.
19. K. G. Vamvoudakis and F. L. Lewis, “Multi-player non-zero-sum games: Online adaptive learning solution of coupled hamilton–jacobi equations,” *Automatica*, vol. 47, no. 8, pp. 1556–1569, 2011.
20. K. G. Vamvoudakis, F. L. Lewis, and G. R. Hudas, “Multi-agent differential graphical games: Online adaptive learning solution for synchronization with optimality,” *Automatica*, vol. 48, no. 8, pp. 1598–1611, 2012.

21. Q. Jiao, H. Modares, S. Xu, F. L. Lewis, and K. G. Vamvoudakis, “Multi-agent zero-sum differential graphical games for disturbance rejection in distributed control,” *Automatica*, vol. 69, pp. 24–34, 2016.
22. T. Bian and Z.-P. Jiang, “Value iteration and adaptive dynamic programming for data-driven adaptive optimal control design,” *Automatica*, vol. 71, pp. 348–360, 2016.
23. D. Vrabie, O. Pastravanu, M. Abu-Khalaf, and F. L. Lewis, “Adaptive optimal control for continuous-time linear systems based on policy iteration,” *Automatica*, vol. 45, no. 2, pp. 477–484, 2009.
24. D. Liu, H. Li, and D. Wang, “Neural-network-based zero-sum game for discrete-time nonlinear systems via iterative adaptive dynamic programming algorithm,” *Neurocomputing*, vol. 110, pp. 92–100, 2013.
25. D. Vrabie and F. Lewis, “Adaptive dynamic programming for online solution of a zero-sum differential game,” *Journal of Control Theory and Applications*, vol. 9, no. 3, pp. 353–360, 2011.
26. F. L. Lewis and D. Vrabie, “Reinforcement learning and adaptive dynamic programming for feedback control,” *IEEE circuits and systems magazine*, vol. 9, no. 3, 2009.
27. R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press Cambridge, 1998, vol. 1, no. 1.
28. F. A. Yaghmaie and D. J. Braun, “Reinforcement learning for a class of continuous-time input constrained optimal control problems,” *Automatica*, vol. 99, pp. 221–227, 2019.
29. A. Isidori, *Nonlinear control systems*. Springer Science & Business Media, 2013.
30. J. Huang, *Nonlinear output regulation: Theory and applications*. SIAM, 2004, vol. 8.
31. F. L. Lewis, D. Vrabie, and V. L. Syrmos, *Optimal control*. John Wiley & Sons, 2012.
32. R. A. Hom and C. R. Johnson, *Topics in matrix analysis*, 1991.
33. H. K. Khalil, *Noninear Systems*. Prentice-Hall, New Jersey, 1996, vol. 2, no. 5.
34. F. Adib Yaghmaie, F. L. Lewis, and R. Su, “Output regulation of heterogeneous multi-agent systems: A graphical game approach,” in *American Control Conference (ACC), 2015*, 2015, pp. 2272–2277.