IET Research Journals

ISSN 1751-8644
doi: 0000000000
www.ietdl.org

Optimal Distributed Learning for Disturbance Rejection in Networked Nonlinear Games under Unknown Dynamics

Farzaneh Tatari[1]*, *Kyriakos G. Vamvoudakis*[2], *Majid Mazouchi*[3]

[1] *Electrical Engineering Department, Semnan University, Semnan, Iran*
[2] *Daniel Guggenheim School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA*
[3] *Electrical Engineering Department, Ferdowsi University of Mashhad, Mashhad, Iran*
* *E-mail: ftatari@ semnan.ac.ir*

Abstract: In this paper, an online distributed optimal adaptive algorithm is introduced for continuous-time nonlinear differential graphical games under unknown systems subject to external disturbances. The proposed algorithm learns online the approximate solution to the coupled Hamilton-Jacobi-Isaacs (HJI) equations. Each of the players in the game uses an actor-critic-disturbance network structure, and an intelligent identifier to find the unknown parameters of the systems. We use recorded past observations concurrently with current data to speed up convergence by exploring the state space. The closed-loop stability and the convergence of the policies to a Nash equilibrium are ensured by using Lyapunov stability theory. Finally, a simulation example shows the efficiency of the proposed algorithm.

1 Introduction

Distributed control of multi-agent systems (MASs) on communication graphs has attracted great attention, motivated by its possible applications in many engineering systems that involve networks. Considerable literature has been developed on distributed control methods to solve consensus problem [1-6]. This problem is mainly separated into two categories: leaderless consensus problem and leader-follower problem. In the second one, which is the problem of interest in this paper, the objective is to design a local control protocol for each agent, which depends only on local information, to ensure that all agents follow the trajectory of an agent called as leader.

It has been recognized in the literature [7-12] that game theory provides a proper framework to study multi-agent problems. Based on differential game theory, the differential graphical game concept is introduced in [7] to provide a framework to solve leader-follower problem in an optimal manner where the tracking error dynamics, actions, and performance index of each follower agent depend on local neighbor information.

The solution of differential graphical games considering the existence of unknown external disturbances is an important issue. However, unknown external disturbances exist in many practical MASs, which are inevitable, and can be a principal cause of poor performance or even worse instability. It is known that solving differential graphical game with external disturbances for the nonlinear systems relies on finding the Nash equilibrium solutions to coupled Hamilton-Jacobi-Isaacs (HJI) equations. However, coupled HJI equations are nonlinear partial differential equations (PDEs) and are difficult or impossible to solve, and may not have global analytical solutions even in simple cases. Therefore, numerical methods are required in order to approximately solve them.

Reinforcement learning (RL) techniques [13] have been employed to solve optimal control problem problems with and without disturbances and modeling uncertainties [14-16]. These techniques have also emerged as an efficient tool to approximately solve the coupled HJI equations online [9,17,18]. In [9], a policy iteration (PI) algorithm is provided to find the solution of coupled HJI equations, but it is limited to linear systems and closed-loop system stability of the equilibrium point is not provided. In [17,18], the authors presented an RL method to design robust adaptive tracking control laws for multi-wheeled mobile robots. But they rely on either complete knowledge of the systems dynamics [9], or at least partial knowledge of the systems dynamics [17,18]. However, most of the practical systems are difficult to model exactly. Furthermore, it is well known that nonlinearities commonly exist in physical systems, and many of physical systems possess higher-order dynamics. Therefore, finding the solution of the coupled HJI equations of nonlinear systems with higher-order unknown dynamics is an important issue from practical point of view, and is also challenging due to the dependency of coupled HJI equations on the communication graphs.

To the best of our knowledge, there has not been any results on differential graphical games of nonlinear systems with higher-order dynamics in the presence of disturbance and completely unknown dynamics. This motivates our research.

Contributions: The contributions of the present paper are threefold. We formulate the problem of nonlinear leader-follower consensus in the presence of disturbances as multi-agent zero-sum differential graphical games under completely unknown nonlinear dynamics. An optimal distributed learning algorithm is proposed to approximately solve the problem of multi-agent zero-sum differential graphical games of general affine nonlinear systems in the presence of external disturbances under unknown dynamics. To this end, the completely unknown nonlinear dynamics are identified online through learning-based identifiers while also using an experience replay technique. Finally, rigorous proofs provide guarantees for convergence of the policies to the approximate Nash equilibrium while guaranteeing closed-loop stability.

Background on graphs: The communication network is described by a graph $\mathrm{Gr} = (V, \Sigma)$, where $V = \{1, 2, \cdots, N\}$ is the set of vertices representing N agents and $\Sigma \subset V \times V$ is the set of edges of the graph. $(i, j) \in \Sigma$, shows that there is an edge from node i to node j. An adjacency matrix $E = \left[e_{ij}\right] \in \Re^{N\times N}$ is often used to represent the graph topology where $e_{ij} = 1$ if $(i, j) \in \Sigma$ and $e_{ij} = 0$ otherwise. The set of neighbors of a node i is $N_i = \{j : (j, i) \in \Sigma\}$ and $i_{N_i} = \{j : (i, j) \in \Sigma\}$ indicates the set of nodes which node i is in their neighborhood. $d_i = \sum_{j\in N_i} e_{ij}$ is the weighted in-degree of node i. The leader is represented by 0 and information is sent from the leader to the agents for which the leader is in their neighborhood.

Structure: The paper is organized as follows. The problem formulation is explained in Section 2 and coupled HJI equations are derived in Section 3. Section 4 explains the MAS approximation-based identifiers. The proposed distributed optimal adaptive learning algorithm in the presence of disturbance and unknown dynamics is introduced in Section 5. The simulation results are discussed in Section 6 and the conclusions are drawn in Section 7.

2 Problem Formulation

Consider the dynamics of each agent, as physical components in a directed strongly connected communication graph (cyber component) to be,

$$\dot{x}_i = f_i(x_i) + g_i(x_i)u_i + k_i(x_i)\omega_i,\ \forall i \in \{1, \cdots, N\}, \quad (1)$$

where $x_i(t) \in \Re^n$ is the measurable state vector, $u_i(t) \in \Re^m$ is the control input, $\omega_i(t) \in \Re^q$ is the external disturbance input, $f_i(x_i) \in \Re^n$, $g_i(x_i) \in \Re^{n\times m}$ and $k_i(x_i) \in \Re^{n\times q}$, $i = 1, \cdots, N$ are respectively the drift, the input and the disturbance dynamics that will be considered unknown in our developments. It is assumed that the closed-loop system $f_i(x_i) + g_i(x_i)u_i + k_i(x_i)\omega_i$, $i = 1, \cdots, N$ is locally Lipschitz (a classical assumption to have a unique solution for any initial condition $x_i(0)$). Consider the uncontrolled leader dynamics that generate the target state as,

$$\dot{x}_0 = f_0(x_0). \quad (2)$$

The local neighborhood tracking error for every agent can be defined as,

$$\delta_i = \sum_{j\in N_i} e_{ij}(x_i - x_j) + e_{i0}(x_i - x_0), \quad (3)$$

where the pinning gain e_{i0} is nonzero for at least one agent which communicates directly with the leader agent and $e_{i0} = 0$ otherwise. The time derivative of (3) is given by,

$$\begin{aligned}\dot{\delta}_i = &\sum_{j\in N_i} e_{ij}(f_i(x_i) - f_j(x_j)) + e_{i0}(f_i(x_i) - f_0(x_0))\\ &+ (d_i + e_{i0})g_i(x_i)u_i - \sum_{j\in N_i} e_{ij}g_j(x_j)u_j\\ &+ (d_i + e_{i0})k_i(x_i)\omega_i - \sum_{j\in N_i} e_{ij}k_j(x_j)\omega_j. \quad (4)\end{aligned}$$

In order to achieve synchronization, a distributed control shall be designed which can keep the tracking error (3) L_2-bounded for $\omega_i(t) \neq 0$, under the unknown dynamics of the MAS.

Bounded L_2-gain synchronization problem. Consider system (4) with measured outputs $y_i = C_i\delta_i$ (where C_i is left invertible and δ_i can be directly measured), disturbances with $\omega_i^{N_i}(t) = \left[\omega_i{}^{\mathrm{T}}(t), \omega_{N_i}{}^{\mathrm{T}}(t)\right]^{\mathrm{T}}$ and performance outputs $z_i(t) = \left[\delta_i{}^{\mathrm{T}}(t), u_i{}^{\mathrm{T}}(t), u_{N_i}{}^{\mathrm{T}}(t)\right]^{\mathrm{T}}$ with $u_{N_i} = \{u_j | j \in N_i\}$, $\omega_{N_i} = \{\omega_j | j \in N_i\}$. It is desired to design control $u_i(t)$ to solve the synchronization problem when $\omega_i(t) = 0$ and also to satisfy the following bounded L_2-gain condition (disturbance attenuation level) for a given $\gamma > \gamma^\star$ when $\omega_i(t) \neq 0$ for all agents,

$$\begin{aligned}\int_0^{\mathrm{T}} \|z_i(t)\|^2 \mathrm{d}t = &\int_0^{\mathrm{T}} \left(Q_i(\delta_i) + u_i{}^{\mathrm{T}}R_{ii}u_i + \sum_{j\in N_i} u_j{}^{\mathrm{T}}R_{ij}u_j\right)\mathrm{d}t\\ \leqslant &\gamma^2 \int_0^{\mathrm{T}} \left\|\omega_i^{N_i}\right\|^2 \mathrm{d}t + \beta(\delta_i(0))\\ = &\gamma^2 \int_0^{\mathrm{T}} \left(\omega_i{}^{\mathrm{T}}T_{ii}\omega_i + \sum_{j\in N_i} \omega_j{}^{\mathrm{T}}T_{ij}\omega_j\right)\mathrm{d}t\\ &+ \beta(\delta_i(0))\end{aligned}$$

for a bounded function β such that $\beta(0) = 0$ [19], with $Q_i(\delta_i) \geq 0$, and the weighting matrices $R_{ii} > 0$, $R_{ij} > 0$, $T_{ii} > 0$ and $T_{ij} > 0$ are symmetric and constant. Let $\gamma^\star$ be the minimum value of γ for which the above disturbance attenuation condition is satisfied.

The local performance index for every agent i is defined as,

$$\begin{aligned}&J_i\left(\delta_i(t), u_i, u_{N_i}, \omega_i, \omega_{N_i}\right) =\\ &\frac{1}{2}\int_t^\infty (Q_i(\delta_i) + u_i{}^{\mathrm{T}}R_{ii}u_i + \sum_{j\in N_i} u_j{}^{\mathrm{T}}R_{ij}u_j\\ &-\gamma^2\omega_i{}^{\mathrm{T}}T_{ii}\omega_i - \gamma^2 \sum_{j\in N_i} \omega_j{}^{\mathrm{T}}T_{ij}\omega_j)\mathrm{d}\tau. \quad (5)\end{aligned}$$

It is shown in [9] that the solution of bounded L_2-gain synchronization problem is equivalent to the solution of the following multi-player zero-sum differential graphical game,

$$V_i{}^\star(\delta_i(t)) = \min_{u_i}\ \max_{\omega_i}\ J_i(\delta_i(t), u_i, u^\star_{N_i}, \omega_i, \omega^\star_{N_i})$$

where the control and disturbance players try to minimize and maximize the value respectively. The game has a unique saddle point solution $(u_i{}^\star, \omega_i{}^\star)$ for every agent if [20],

$$\begin{aligned}V_i{}^\star(\delta_i(t)) &= \min_{u_i}\max_{\omega_i} J_i(\delta_i(t), u_i, u^\star_{N_i}, \omega_i, \omega^\star_{N_i})\\ &= \max_{\omega_i}\ \min_{u_i} J_i(\delta_i(t), u_i, u^\star_{N_i}, \omega_i, \omega^\star_{N_i}).\end{aligned}$$

The associated value $V_i^\star$ is the value of the game. This is equivalent to the following Nash equilibrium condition,

$$\begin{aligned}J_i(u_i^\star, u^\star_{N_i}, \omega_i, \omega^\star_{N_i}) &\leqslant J_i(u_i^\star, u^\star_{N_i}, \omega_i^\star, \omega^\star_{N_i})\\ &\leqslant J_i(u_i, u^\star_{N_i}, \omega_i^\star, \omega^\star_{N_i}), \forall u_i, \omega_i.\end{aligned}$$

Therefore, given (4) the value function for every node i is given $\forall t$ as,

$$\begin{aligned}V_i{}^\star(\delta_i(t)) = &\frac{1}{2}\min_{u_i}\max_{\omega_i}\int_t^\infty (Q_i(\delta_i) + u_i{}^{\mathrm{T}}R_{ii}u_i\\ &+ \sum_{j\in N_i} u_j^{\star\mathrm{T}}R_{ij}u_j^\star - \gamma^2\omega_i{}^{\mathrm{T}}T_{ii}\omega_i\\ &- \gamma^2 \sum_{j\in N_i} \omega_j^{\star\mathrm{T}}T_{ij}\omega_j^\star)\mathrm{d}\tau. \quad (6)\end{aligned}$$

Remark 1. The inclusion of a game-theoretic control framework to the learning setting guarantees a high degree of robustness which is required to maintain a sufficient stability margin of the closed-loop system in terms of parametric uncertainties and output disturbances. □

3 Coupled HJI Equations

The value function (6) can be equivalently described by the following Lyapunov equation in terms of the Hamiltonian function,

$$\begin{aligned}&H_i(\delta_i, \nabla V_i, u_i, u_{N_i}, \omega_i, \omega_{N_i}) \equiv \frac{1}{2}Q_i(\delta_i) + \nabla V_i{}^{\mathrm{T}} \times\\ &[\sum_{j\in N_i} e_{ij}(f_i(x_i) - f_j(x_j)) + e_{i0}(f_i(x_i) - f_0(x_0))\\ &+ (d_i + e_{i0})(g_i(x_i)u_i + k_i(x_i)\omega_i)\\ &- \sum_{j\in N_i} e_{ij}g_j(x_j)u_j - \sum_{j\in N_i} e_{ij}k_j(x_j)\omega_j]\\ &+\frac{1}{2}u_i{}^{\mathrm{T}}R_{ii}u_i + \frac{1}{2}\sum_{j\in N_i} u_j{}^{\mathrm{T}}R_{ij}u_j\\ &-\frac{1}{2}\gamma^2\omega_i{}^{\mathrm{T}}T_{ii}\omega_i - \frac{1}{2}\gamma^2 \sum_{j\in N_i} \omega_j{}^{\mathrm{T}}T_{ij}\omega_j = 0, \quad (7)\end{aligned}$$

IET Research Journals, pp. 1–11

where $\nabla V_i = \frac{\partial V_i}{\partial \delta_i} \in \Re^n$ and $V_i(0) = 0$.

After employing the stationarity conditions in the Hamiltonians, one has,

$$\frac{\partial H_i}{\partial u_i} = 0 \rightarrow {u_i}^{\star} = -(d_i + e_{i0}) R_{ii}^{-1} {g_i}^{\mathrm{T}}(x_i) \nabla V_i, \quad (8)$$

$$\frac{\partial H_i}{\partial \omega_i} = 0 \rightarrow {\omega_i}^{\star} = \frac{1}{\gamma^2}(d_i + e_{i0}) T_{ii}^{-1} {k_i}^{\mathrm{T}}(x_i) \nabla V_i. \quad (9)$$

Substituting (8) and (9), in (7), yields the following coupled HJI equations,

$$\begin{aligned}
&{\nabla V_i}^{\mathrm{T}}[\sum_{j \in N_i} e_{ij}(f_i(x_i) - f_j(x_j)) \\
&+ e_{i0}(f_i(x_i) - f_0(x_0)) - {c_i}^2 g_i(x_i) R_{ii}^{-1} {g_i}^{\mathrm{T}}(x_i) \nabla V_i \\
&+ \sum_{j \in N_i} e_{ij} c_j g_j(x_j) R_{jj}^{-1} \times {g_j}^{\mathrm{T}}(x_j) \nabla V_j \\
&- \frac{1}{\gamma^2} \sum_{j \in N_i} e_{ij} c_j k_j(x_j) T_{jj}^{-1} {k_j}^{\mathrm{T}}(x_j) \nabla V_j \\
&+ \frac{1}{\gamma^2} {c_i}^2 k_i(x_i) T_{ii}^{-1} {k_i}^{\mathrm{T}}(x_i) \nabla V_i] + \frac{1}{2} Q_i(\delta_i) \\
&+ \frac{1}{2} \sum_{j \in N_i} {c_j}^2 {\nabla V_j}^{\mathrm{T}} g_j(x_j) R_{jj}^{-1} R_{ij} R_{jj}^{-1} {g_j}^{\mathrm{T}}(x_j) \nabla V_j \\
&- \frac{1}{2\gamma^2} \sum_{j \in N_i} {c_j}^2 {\nabla V_j}^{\mathrm{T}} k_j(x_j) T_{jj}^{-1} T_{ij} T_{jj}^{-1} {k_j}^{\mathrm{T}}(x_j) \nabla V_j \\
&- \frac{1}{2\gamma^2} {c_i}^2 {\nabla V_i}^{\mathrm{T}} k_i(x_i) T_{ii}^{-1} {k_i}^{\mathrm{T}}(x_i) \nabla V_i \\
&+ \frac{1}{2} {c_i}^2 {\nabla V_i}^{\mathrm{T}} g_i(x_i) R_{ii}^{-1} {g_i}^{\mathrm{T}}(x_i) \nabla V_i = 0, \quad (10)
\end{aligned}$$

where $c_i := d_i + e_{i0}$, and with a boundary condition $V_i(0) = 0$. For a given solution V_i to (10), after defining ${u_i}^{\star} = u_i(V_i)$, ${\omega_i}^{\star} = \omega_i(V_i)$ in terms of V_i, we can rewrite (10) as $H_i(\delta_i, \nabla V_i, u_i^{\star}, u_{N_i}^{\star}, \omega_i^{\star}, \omega_{N_i}^{\star}) = 0,\ V_i(0) = 0$. The coupled HJI equations (10) are highly nonlinear partial differential equations and require the complete knowledge of the dynamics which make these equations difficult to solve. For that reason we will use approximation-based techniques.

Remark 2. In system (4) with the corresponding value function (6), the optimal control policy and the worst case disturbance, minimize and maximize respectively the cost function (6). Therefore, the optimal control policy and the worst-case disturbance can be obtained by employing the stationarity conditions (8) and (9) respectively. For every agent i, as it is shown in (8) and (9), the optimal control policy ${u_i}^{\star}$ and the worst case disturbance ${\omega_i}^{\star}$ are both functions of the local tracking error δ_i for the agent i (due to the term $\nabla V_i = \frac{\partial V_i}{\partial \delta_i}$). Hence ${u_i}^{\star}$ and ${\omega_i}^{\star}$ are both distributed optimal control, and worst case disturbance policies. □

4 Approximation-Based System Identification

Before we proceed, the following definition and assumptions are needed.

Definition 1 (Persistence of Excitation (PE)) The bounded vector signal $\bar{\Theta}_i(t),\ i = 1, \cdots, N$ is PE over the interval $[t, t + T_i]$ if there exists $T_i > 0$, $\gamma_i > 0$ and $\gamma_{i+N} > 0$ such that for all t,

$$\gamma_i I \leqslant \int_t^{t+T_i} \bar{\Theta}_i(\tau) \bar{\Theta}_i^{\mathrm{T}}(\tau) \mathrm{d}\tau \leqslant \gamma_{i+N} I\ ; i = 1, \ldots, N.$$

□

Assumption 1. Given admissible feedback control policies, then the nonlinear Lyapunov equations (7) have locally smooth solutions $V_i^{\star}(\delta_i)$. □

Remark 3. Assumption 1 is widely used, since optimal control problems do not necessarily have smooth or even continuous value functions [17]. In this paper all derivations are performed under the assumption of smooth solutions to (7) and (10) (see [7,8,11]). This will allow us to use the Weierstrass high-order approximation theorem [22]. □

Assumption 2. For a given compact set $\Omega \subset \Re^n$ and $i = 1, \ldots, N$, the reconstruction errors, the approximator basis functions, and the gradients of both are bounded. □

Remark 4. Assumption 2 is standard in the literature [7,8] according to Weierstrass high-order approximation theorem. Note further that, the approximators used are the so-called functional link neural networks (see [21] for more details), for which the activation functions σ_i for $i = 1, \cdots, N$ can be some squashing functions, such as the standard sigmoid, Gaussian, and hyperbolic tangent functions. Furthermore, the bounds mentioned above are only used for the stability analysis, and they are actually not used in the controller design. □

Motivated by [23], in order to identify the unknown dynamics of every agent i, $i = 1, \cdots, N$ in a compact set Ω, we will use identifiers as follows,

$$\begin{aligned}
f_i(x_i) &= {\theta_i}^{\star} \xi_i(x_i) + \varepsilon_{f_i}, \\
g_i(x_i) &= {\psi_i}^{\star} \varsigma_i(x_i) + \varepsilon_{g_i}, \quad (11) \\
k_i(x_i) &= {\beta_i}^{\star} \vartheta_i(x_i) + \varepsilon_{k_i},
\end{aligned}$$

where $\theta_i^{\star} \in \Re^{n \times k_{\theta_i}}$, $\psi_i^{\star} \in \Re^{n \times k_{\psi_i}}$, $\beta_i^{\star} \in \Re^{n \times k_{\beta_i}}$ are unknown weights, $\xi_i \in \Re^{k_{\theta_i}}$, $\varsigma_i \in \Re^{k_{\psi_i} \times m}$, $\vartheta_i \in \Re^{k_{\beta_i} \times q}$ are basis functions, ε_{f_i}, ε_{g_i} and ε_{k_i} are the reconstruction errors. By using (11), the system (1) can be re-written as

$$\dot{x}_i = \varphi_i^{\star} z(x_i, u_i, \omega_i) + \varepsilon_i,\ \forall i, \quad (12)$$

where $z(x_i, u_i, \omega_i) = [{\xi_i}^{\mathrm{T}},\ {u_i}^{\mathrm{T}} {\varsigma_i}^{\mathrm{T}}, {\omega_i}^{\mathrm{T}} {\vartheta_i}^{\mathrm{T}}]^{\mathrm{T}}$ is the regressor vector, $\varphi_i^{\star} = [\theta_i^{\star}, \psi_i^{\star}, \beta_i^{\star}]$ and $\varepsilon_i = \varepsilon_{f_i} + \varepsilon_{g_i} + \varepsilon_{k_i}$. Using Assumption 1 we have $\|\varepsilon_i\| \leqslant \bar{\varepsilon}_i$, where $\bar{\varepsilon}_i = \bar{\varepsilon}_{f_i} + \bar{\varepsilon}_{g_i} + \bar{\varepsilon}_{k_i}$ and $\|\varepsilon_{f_i}\| \leqslant \bar{\varepsilon}_{f_i}$, $\|\varepsilon_{g_i}\| \leqslant \bar{\varepsilon}_{g_i}$, $\|\varepsilon_{k_i}\| \leqslant \bar{\varepsilon}_{k_i}$.

The dynamics (12) can be written as,

$$\dot{x}_i = -Ax_i + \varphi_i^{\star} z(x_i, u_i, \omega_i) + Ax_i + \varepsilon_i$$

where $A = aI_{n \times n},\ a > 0,\ i = 1, \cdots, N$. The following lemma adopted from [23] provides a filtered regressor for (12).

Lemma 1. *The solution of (12) can be expressed as*

$$x_i = {\varphi_i}^{\star} h_i(x_i) + a l_i(x_i) + \varepsilon_{x_i}, \quad (13)$$

$$\begin{aligned}
\dot{h}_i(x_i) &= -a h_i(x_i) + z(x_i, u_i, \omega_i); \\
\dot{l}_i(x_i) &= -A l_i(x_i) + x_i
\end{aligned} \quad (14)$$

where $h_i(x_i) = \int_0^t e^{-a(t-\tau)} z(x_i(\tau), u_i(\tau), \omega_i(\tau)) d\tau$, $h_i(x_i) \in \Re^{k_{\theta_i} + k_{\psi_i} + k_{\beta_i}}$ *is the filtered regressor version of* $z(x_i, u_i, u_i)$, $\varepsilon_{x_i} = e^{-At} x_i(0) + \int_0^t e^{-A(t-\tau)} \varepsilon_i d\tau$, $l_i(x_i) = \int_0^t e^{-A(t-\tau)} x_i(\tau)$ $d\tau$, *and* $x_i(0)$ *is the initial state of (12).*

Each side of (13) is divided with a normalizing signal $n_{s_i} = 1 + {h_i}^{\mathrm{T}} h_i + {l_i}^{\mathrm{T}} l_i$ to obtain,

$$\begin{aligned}
\bar{x}_i &= {\varphi_i}^{\star} \bar{h}_i(x_i) + a \bar{l}_i(x_i) + \bar{\varepsilon}_{x_i}, \\
\bar{x}_i &= \frac{x_i}{n_{s_i}},\ \bar{h}_i = \frac{h_i}{n_{s_i}},\ \bar{l}_i = \frac{l_i}{n_{s_i}},\ \bar{\varepsilon}_{x_i} = \frac{\varepsilon_{x_i}}{n_{s_i}}.
\end{aligned} \quad (15)$$

Based on Lemma 1 and equation (15), we consider the i^{th} identifier weights estimator to be of the form,

$$\dot{\hat{\bar{x}}}_i = \hat{\varphi}_i \bar{h}_i(x_i) + a\bar{l}_i(x_i),\ i = 1, \cdots, N,$$

where $\hat{\varphi}_i = [\hat{\theta}_i, \hat{\psi}_i, \hat{\beta}_i] \in \Re^{n\times(k_{\theta_i}+k_{\psi_i}+k_{\beta_i})}$ is the estimated value of weights matrix ${\varphi_i}^{\star}$ at time t, for agent i. We shall define the state estimation error of agent i, $i = 1, \cdots, N$ as

$$\begin{aligned} e_i(t) &= \hat{\bar{x}}_i - \bar{x}_i = \tilde{\varphi}_i(t)\bar{h}_i(x_i(t)) - \bar{\varepsilon}_{x_i}\ , \\ \tilde{\varphi}_i(t) &= \hat{\varphi}_i(t) - {\varphi_i}^{\star}(t), \end{aligned} \tag{16}$$

where $\tilde{\varphi}_i(t) = [\tilde{\theta}_i\ , \tilde{\psi}_i, \tilde{\beta}_i]$ is the parameter estimation error and $\tilde{\theta}_i = \hat{\theta}_i - {\theta_i}^{\star},\ \tilde{\psi}_i = \hat{\psi}_i - {\psi_i}^{\star},\ \tilde{\beta}_i = \hat{\beta}_i - {\beta_i}^{\star}$. We shall use the idea of experience replay [23] which employs recorded observations along with current data to obtain the tuning law of the identifier weights.

Define the recorded past data that is collected and stored in the history stack of each agent i, $i = 1, \cdots, N$ at times $t_1, \cdots, t_{p_i}$ as,

$$Z_i = [\bar{h}_i(x_i(t_1)), \cdots, \bar{h}_i(x_i(t_{p_i}))].$$

Consider now p_i as the number of data points stored in the history stack of agent i as Z_i which must contain as many linearly independent elements as the dimension of the basis of the uncertainty $h_i(x_i)$ in (13) (Z_i rank condition) in order to satisfy the PE condition.

The tuning algorithm for the i^{th} agent identifier weights is given as,

$$\begin{aligned} \dot{\hat{\varphi}}_i(t) &= -\Gamma_i \bar{h}_i(x_i(t)){e_i}^{\mathrm{T}}(t) \\ &\quad -\Gamma_i \sum_{k=1}^{p_i} \bar{h}_i(x_i(t_k)){e_i}^{\mathrm{T}}(t_k), i = 1, \cdots, N, \end{aligned} \tag{17}$$

where $\Gamma_i \succ 0,\ i = 1, \cdots, N$ indicates a positive definite learning rate matrix which affects the speed of learning.

Theorem 1. *Consider the system given by (12). Let the online i^{th} identifier tuning law be given by the update law of (17) with a filtered regressor given by (14). Then, given that the recorded data points vector Z_i has full rank condition, for a bounded model approximation error, the identifier weights estimation errors are uniformly ultimately bounded (UUB), i.e., there exists a bound $B_{\tilde{\varphi}_i}$ and time $T(B_{\tilde{\varphi}_i}, \tilde{\varphi}_i(0)) = T_{\tilde{\varphi}_i}$ such that $\|\tilde{\varphi}_i\| \leqslant B_{\tilde{\varphi}_i}$ for all $t \geqslant t_0 + T_{\tilde{\varphi}_i}$.*

Proof: The proof is an extension of the proof in [23]. It can be shown that given that the rank condition is satisfied, the identifier approximation error $\tilde{\varphi}_i$ is bounded outside $\Omega_{\tilde{\varphi}_i}$ area,

$$\Omega_{\tilde{\varphi}_i} = \left\{ \tilde{\varphi}_i : \|\tilde{\varphi}_i\| \leqslant B_{\tilde{\varphi}_i}, B_{\tilde{\varphi}_i} = \frac{\bar{\varepsilon}_{T_i}(p_i+1)}{a\ \sigma_{\min}(H_{\tilde{\varphi}_i})} \right\} \tag{18}$$

where $\sigma_{\min}$ stands for the smallest singular value,

$$\begin{aligned} H_{\tilde{\varphi}_i} &= n_{s_i}(\bar{h}_i(\delta_i(t))^{\mathrm{T}}\bar{h}_i(\delta_i(t)) + \sum_{k=1}^{p_i} \bar{h}_i(\delta_i(t_k))^{\mathrm{T}}\bar{h}_i(\delta_i(t_k))), \\ \varepsilon_{T_i} &= \sum_{j\in N_i} e_{ij}(\varepsilon_{f_i} - \varepsilon_{f_j}) + e_{i0}(\varepsilon_{f_i} - \varepsilon_{f_0}) + (d_i + e_{i0})\varepsilon_{g_i}u_i \\ &\quad - \sum_{j\in N_i} e_{ij}\varepsilon_{g_j}u_j + (d_i + e_{i0})\varepsilon_{k_i}\omega_i - \sum_{j\in N_i} e_{ij}\varepsilon_{k_j}\omega_j, \end{aligned}$$

and finally we have $\|\varepsilon_{T_i}\| \leqslant \bar{\varepsilon}_{T_i}, \bar{\varepsilon}_{T_i} = \sum_{j\in N_i} e_{ij} \times (\bar{\varepsilon}_{f_i} + \bar{\varepsilon}_{f_j}) + e_{i0}(\bar{\varepsilon}_{f_i} + \bar{\varepsilon}_{f_0}) + (d_i + e_{i0})\bar{\varepsilon}_{g_i}\|u_i\| + \sum_{j\in N_i} e_{ij} \times \bar{\varepsilon}_{g_j}\|u_j\| + (d_i + e_{i0})\bar{\varepsilon}_{k_i}\|\omega_i\| + \sum_{j\in N_i} e_{ij}\bar{\varepsilon}_{k_j}\|\omega_j\|$. ∎

Remark 5. In order to minimize $B_{\tilde{\varphi}_i}$, a and p_i must be chosen appropriately. One can decrease $B_{\tilde{\varphi}_i}$ by choosing a large design parameter a and the number of recorded data points p_i shall maximize $\sigma_{\min}(H_{\tilde{\varphi}_i})$ to reduce the error bound. □

Now, (4) can be written in a compact form as,

$$\dot{\delta}_i = {\varphi_{\delta_i}}^{\star} z_i(x_i, x_j, u_i, u_{N_i}, \omega_i, \omega_{N_i}) + \varepsilon_{T_i}, \tag{19}$$

where for $i = 1, \cdots, N,\ j \in N_i$,

$$\begin{aligned} & z_i(x_i, x_j, u_i, u_{N_i}, \omega_i, \omega_{N_i}) = [z^{\mathrm{T}}_{e_{ij}\xi_i}, z^{\mathrm{T}}_{-e_{ij}\xi_j}, e_{i0}{\xi_i}^{\mathrm{T}}, \\ & -e_{i0}{\xi_0}^{\mathrm{T}}, (d_i + e_{i0})(\varsigma_i u_i)^{\mathrm{T}}, z^{\mathrm{T}}_{-e_{ij}\xi_j u_j}, (d_i + e_{i0})(\vartheta_i\omega_i)^{\mathrm{T}} \\ & \qquad , z^{\mathrm{T}}_{-e_{ij}\vartheta_j\omega_j}]^{\mathrm{T}}, \\ & z_{e_{ij}\xi_i} = \left\{ e_{ij}\xi_i | j \in N_i \right\}, z_{-e_{ij}\xi_j} = \left\{ -e_{ij}\xi_j | j \in N_i \right\}, \\ & z_{-e_{ij}\varsigma_j u_j} = \left\{ -e_{ij}\varsigma_j u_j | j \in N_i \right\}, \\ & z_{-e_{ij}\vartheta_j\omega_j} = \left\{ -e_{ij}\vartheta_j\omega_j | j \in N_i \right\}, \\ & {\varphi_{\theta_i}}^{\star} = [{\theta_i}^{\star} \cdots {\theta_i}^{\star}], {\varphi_{\theta_j}}^{\star} = [{\theta_j}^{\star} \cdots {\theta_j}^{\star}], \\ & {\varphi_{\psi_j}}^{\star} = [{\psi_j}^{\star} \cdots {\psi_j}^{\star}], {\varphi_{\beta_j}}^{\star} = [{\beta_j}^{\star} \cdots {\beta_j}^{\star}], \\ & {\varphi_i}^{\star} = [{\varphi_{\theta_i}}^{\star}, {\varphi_{\theta_j}}^{\star}, {\theta_i}^{\star}, {\theta_0}^{\star}, {\psi_i}^{\star}, {\varphi_{\psi_j}}^{\star}, {\beta_i}^{\star}, {\varphi_{\beta_j}}^{\star}], \\ & \mathrm{Card}({\varphi_{\theta_i}}^{\star}) = \mathrm{Card}({\varphi_{\theta_j}}^{\star}) = \mathrm{Card}({\varphi_{\psi_j}}^{\star}) = N_i, \end{aligned}$$

where $\mathrm{Card}(\cdot)$ is the cardinality measure.

Therefore the local error dynamics (4) is approximated as

$$\dot{\delta}_i = \hat{\varphi}_{\delta_i} z_i(x_i, x_j, u_i, u_{N_i}, \omega_i, \omega_{N_i}) + \varepsilon_{T_i}, \tag{20}$$

where $\hat{\varphi}_{\delta_i} = [\hat{\varphi}_{\theta_i}, \hat{\varphi}_{\theta_j}, \hat{\theta}_i, \hat{\theta}_0, \hat{\psi}_i, \hat{\varphi}_{\psi_j}, \hat{\beta}_i, \hat{\varphi}_{\beta_j}] \in \Re^{n\times d^i}$, are the estimated values of ${\varphi_i}^{\star}$ with $\hat{\varphi}_{\theta_i} = [\hat{\theta}_i, \cdots, \hat{\theta}_i]$, $\hat{\varphi}_{\theta_j} = \left\{ \hat{\theta}_j | j \in N_i \right\}, \hat{\varphi}_{\psi_j} = \left\{ \hat{\psi}_j | j \in N_i \right\},\ \hat{\varphi}_{\beta_j} = \left\{ \hat{\beta}_j | j \in N_i \right\}$.

It is worth noting that $\tilde{\varphi}_{\delta_i}(t) = \hat{\varphi}_{\delta_i}(t) - {\varphi_{\delta_i}}^{\star}(t)$ is UUB based on Theorem 1 (i.e. $\|\tilde{\varphi}_{\delta_i}\| \leqslant b_{\tilde{\varphi}_i}$).

Remark 6. In RL, there exist some methods which are model-free, and system identification is not required. However, due to the coupling terms in the coupled HJI equations (10) and their dependence on graph topology and unknown dynamics, the model-free solution in [24] cannot be straightforwardly extended to solve the existing coupled HJI equations. To overcome the difficulty of solving the coupled HJI equations for MASs under unknown dynamics, this paper proposes to use a simple system identifier along with a learning algorithm for every agent to approximately solve the coupled HJI equation and identify the unknown dynamics simultaneously. □

5 Learning Algorithm

We will now use, actor, critic and worst-case disturbance approximators to solve the coupled HJI equations (10). The critic will approximate the cost of each agent, and two actors will be used to approximate the optimal control and the worst-case disturbance.

5.1 Critic Approximators

According to the Weierstrass higher-order approximation theorem [22], there exist independent basis sets $\sigma_i(\delta_i) : \Omega \to \Re^{K_i}$ such that $\sigma_i(0) = 0$, $\nabla\sigma_i(0) = 0$ and constant approximator weights $W_i \in \Re^{K_i},\ i = 1, \cdots, N$ such that the solutions V_i and $\nabla V_i = \frac{\partial V_i}{\partial \delta_i}$ are uniformly approximated on a compact set Ω as follows,

$$V_i = {W_i}^{\mathrm{T}} \sigma_i(\delta_i) + v_i(\delta_i), \ \ i = 1, \cdots, N, \tag{21}$$

$$\nabla V_i = {\nabla \sigma_i}^{\mathrm{T}} W_i + \nabla v_i, \ \ i = 1, \cdots, N, \tag{22}$$

where $\sigma_i(\delta_i) \in \Re^{K_i}$ are the activation function vectors, $K_i, i = 1, \cdots, N$ is the number of basis functions and $v_i(\delta_i)$ are the residuals.

Remark 7. The approximators (11) and (21) are functional link approximators in a Fourier series form which can approximate every function and its derivative. □

The approximation errors $v_i \to 0$ and $\nabla v_i \to 0$ uniformly, as $K_i \to \infty$. Note that according to Assumption 1, we also have $\|v_i\| \leqslant b_{v_i}$, $\|\nabla v_i\| \leqslant b_{\nabla v_i}$, $\|\sigma_i\| \leqslant b_{\sigma_i}$ and $\|\nabla \sigma_i\| \leqslant b_{\nabla \sigma_i}$, $\forall i$. By using critic approximators (21), and fixed feedback policies u_i and u_{N_i} the Hamiltonians (7) can be approximated as follows,

$$\begin{aligned} H_i(\delta_i, W_i, &u_i, u_{N_i}, \omega_i, \omega_{N_i}) = \\ & {W_i}^{\mathrm{T}} \nabla \sigma_i \left[{\varphi_{\delta_i}}^{\star} z_i(\delta_i, \delta_j, u_i, u_{N_i}, \omega_i, \omega_{N_i}) + \varepsilon_{T_i} \right] \\ & + \frac{1}{2} Q_i(\delta_i) + \frac{1}{2} \sum_{j \in N_i} {u_j}^{\mathrm{T}} R_{ij} u_j + \frac{1}{2} {u_i}^{\mathrm{T}} R_{ii} u_i \\ & - \frac{1}{2} \gamma^2 {\omega_i}^{\mathrm{T}} T_{ii} \omega_i - \frac{1}{2} \gamma^2 \sum_{j \in N_i} {\omega_j}^{\mathrm{T}} T_{ij} \omega_j = e_{B_i}. \end{aligned}$$

Note that $e_{B_i} = -(\nabla v_i)^{\mathrm{T}} [{\varphi_{\delta_i}}^{\star} z_i(x_i, x_j, u_i, u_{N_i}, \omega_i, \omega_{N_i}) + \varepsilon_{T_i}]$ and according to Assumption 1, $Sup_{x \in \Omega} \|e_{B_i}\| \leqslant \bar{e}_i$, $i = 1, \cdots, N$ on the compact set Ω.

Assumption 3. For a given compact set $\Omega \subset \Re^n$ and $i = 1, \ldots, N$ we assume that: (i) $\|f_i(x_i)\| \leqslant b_f \|x_i\|$; (ii) $g_i(x_i)$ and $k_i(x_i)$ are bounded by constants $\|g_i(x_i)\| \leqslant b_{g_i}$ and $\|k_i(x_i)\| \leqslant b_{k_i}$ respectively; and (iii) the critic approximators weights are bounded by known constants $\|W_i\| < W_{i\,\max}$. □

Remark 8. Assumption 3 is a standard assumption in neuro-adaptive control literature [8], [21], [23]. Although Assumption 3 restricts the considered class of nonlinear systems, many practical systems (e.g., robotic systems [24] and aircraft systems [25]) satisfy such a property. □

The critic approximators output $\hat{V}_i(\delta_i)$ and the approximate Bellman equations can respectively be written as,

$$\hat{V}_i = \hat{W}_i^{\mathrm{T}} \sigma_i(\delta_i), \tag{23}$$

$$\begin{aligned} e_{H_i} = & \hat{W}_i^{\mathrm{T}} \nabla \sigma_i \hat{\varphi}_{\delta_i} \left[z_i(x_i, x_j, u_i, u_{N_i}, \omega_i, \omega_{N_i}) \right]^{\mathrm{T}} \\ & + \tfrac{1}{2} Q_i(\delta_i) + \tfrac{1}{2} {u_i}^{\mathrm{T}} R_{ii} u_i + \tfrac{1}{2} \sum_{j \in N_i} {u_j}^{\mathrm{T}} R_{ij} u_j \\ & - \tfrac{1}{2} \gamma^2 {\omega_i}^{\mathrm{T}} R_{ii} \omega_i - \tfrac{1}{2} \gamma^2 \sum_{j \in N_i} {\omega_j}^{\mathrm{T}} T_{ij} \omega_j \end{aligned} \tag{24}$$

where $\hat{\varphi}_{\delta_i}$ and $\hat{W}_i \in \Re^{K_i}$ are the current estimated values of ${\varphi_{\delta_i}}^{\star}$ and $W_i \in \Re^{K_i}$ respectively. It is desired to pick $\hat{W}_i$ to minimize the squared residual error, $E_i = \frac{1}{2} {e_{H_i}}^{\mathrm{T}} e_{H_i}$. Hence, the gradient based tuning law for the critic weights of each player is selected as follows,

$$\begin{aligned} \dot{\hat{W}}_i = & -\alpha_i \frac{\partial E_i}{\partial \hat{W}_i} = -\alpha_i e_{H_i} \frac{\partial e_{H_i}}{\partial \hat{W}_i} = -\alpha_i \frac{\bar{B}_i}{m_{s_i}} [\hat{B}_i^{\mathrm{T}} \hat{W}_i \\ & + \frac{1}{2} Q_i(\delta_i) + \frac{1}{2} {u_i}^{\mathrm{T}} R_{ii} u_i + \frac{1}{2} \sum_{j \in N_i} {u_j}^{\mathrm{T}} R_{ij} u_j \\ & - \frac{1}{2} \gamma^2 {\omega_i}^{\mathrm{T}} T_{ii} \omega_i - \frac{1}{2} \gamma^2 \sum_{j \in N_i} \omega^{\mathrm{T}} T_{ij} \omega_j], \end{aligned} \tag{25}$$

where $\hat{B}_i = \nabla \sigma_i(\hat{\varphi}_{\delta_i} \left[z_i(x_i, x_j, u_i, u_{N_i}, \omega_i, \omega_{N_i}) \right])$, $m_{s_i} = 1 + \hat{B}_i^{\mathrm{T}} \hat{B}_i$, $\bar{B}_i = \frac{\hat{B}_i}{1 + \hat{B}_i^{\mathrm{T}} \hat{B}_i}$. $\alpha_i > 0$, $i = 1, \cdots, N$ is the learning rate that determines the speed of convergence.

Lemma 2. *Consider $(u_i, u_{N_i}, \omega_i, \omega_{N_i})$, $i = 1, \cdots, N$ be a set containing given admissible feedback control policies and disturbances, let (25) be the tuning of the critic approximators weights, along with (17) for tuning the identifiers weights and assume that $\bar{B}_i = \frac{\hat{B}_i}{(1 + \hat{B}_i^T \hat{B}_i)}$, $i = 1, \cdots, N$ is PE. Then for bounded reconstruction errors, the critic weights estimation errors converge exponentially to the residual set*

$$\begin{aligned} & \eta_{i_1} e^{-\eta_{i_2} t} + \frac{\alpha_i}{m_{s_i} \eta_{i_2}} b_{\nabla \sigma_i} \|W_i\| \times \\ & (\left\| z_i(\delta_i, \delta_j, u_i, u_j, \omega_i, \omega_j)_{j \in N_i} \right\| b_{\tilde{\varphi}_i} + \bar{\varepsilon}_{T_i}) \\ & + \frac{\alpha_i}{m_{s_i} \eta_{i_2}} \bar{e}_i, \qquad i = 1, \cdots, N \end{aligned}$$

for some $\eta_{i_1}, \eta_{i_2} > 0$.

Proof: From the coupled HJI equations we have,

$$\begin{aligned} -{W_i}^{\mathrm{T}} \nabla \sigma_i & \left[{\varphi_{\delta_i}}^{\star} z_i(\delta_i, \delta_j, u_i, u_{N_i}, \omega_i, \omega_{N_i}) + \varepsilon_{T_i} \right] + e_{B_i} = \\ & + \frac{1}{2} Q_i(\delta_i) + \frac{1}{2} \sum_{j \in N_i} {u_j}^{\mathrm{T}} R_{ij} u_j + \frac{1}{2} {u_i}^{\mathrm{T}} R_{ii} u_i \\ & - \frac{1}{2} \gamma^2 {\omega_i}^{\mathrm{T}} T_{ii} \omega_i - \frac{1}{2} \gamma^2 \sum_{j \in N_i} {\omega_j}^{\mathrm{T}} T_{ij} \omega_j. \end{aligned} \tag{26}$$

Now, substituting (26) in (24) and doing some simple algebraic manipulations, we obtain

$$\begin{aligned} e_{H_i} = & {W_i}^{\mathrm{T}} \nabla \sigma_i (\tilde{\varphi}_{\delta_i} z_i(x_i, x_j, u_i, u_{N_i}, \omega_i, \omega_{N_i}) + \varepsilon_{T_i}) \\ & - \tilde{W}_i^{\mathrm{T}} \nabla \sigma_i (\hat{\varphi}_{\delta_i} z_i(x_i, x_j, u_i, u_{N_i}, \omega_i, \omega_{N_i})) + e_{B_i} \end{aligned} \tag{27}$$

Substituting (27) in (25), yields,

$$\begin{aligned} \dot{\tilde{W}}_i = & -\alpha_i \bar{B}_i \bar{B}_i^{\mathrm{T}} \tilde{W}_i + \alpha_i \frac{\bar{B}_i}{m_{s_i}} ({W_i}^{\mathrm{T}} \nabla \sigma_i \times \\ & (\tilde{\varphi}_{\delta_i} z_i(x_i, x_j, u_i, u_{N_i}, \omega_i, \omega_{N_i}) + \varepsilon_{T_i}) + e_{B_i}) \end{aligned} \tag{28}$$

Assuming that, (28) is a linear time-varying system with an input given by ${W_i}^{\mathrm{T}} \nabla \sigma_i (\tilde{\varphi}_{\delta_i} z_i(x_i, x_j, u_i, u_{N_i}, \omega_i, \omega_{N_i}) + \varepsilon_{T_i}) + e_{B_i}$, $i = 1, \cdots, N$,then, the closed-form solution $\tilde{W}_i$, $i = 1, \cdots, N$ is given as

$$\begin{aligned} \tilde{W}_i(t) = & \phi_i(t, t_0) \, \tilde{W}_i(0) + \alpha_i \int_{t_0}^{\mathrm{T}} \phi_i(\tau, t_0) \, \frac{\bar{B}_i}{m_{s_i}} ({W_i}^{\mathrm{T}}(\tau) \nabla \sigma_i \times \\ & (\tilde{\varphi}_{\delta_i} z_i(x_i, x_j, u_i, u_{N_i}, \omega_i, \omega_{N_i}) + \varepsilon_{T_i}) \\ & + e_{B_i}) \, \mathrm{d}\tau \end{aligned} \tag{29}$$

where the state transition matrix can be found from,

$$\frac{\partial \phi_i(t, t_0)}{\partial t} = -\alpha_i \bar{B}_i \bar{B}_i^{\mathrm{T}} \phi_i(t, t_0), \quad \phi_i(t_0, t_0) = I. \tag{30}$$

The state transition matrix ϕ_i, $i = 1, \cdots, N$ has an exponentially stable equilibrium point provided that $\bar{B}_i$, $i = 1, \cdots, N$ is PE [27]. Using Assumption 1 and the fact that $\bar{B}_i$, $i = 1, \cdots, N$ is PE and

$\left\|\bar{B}_i\right\| \leqslant 1,\ i = 1, \cdots, N$ we obtain

$$\begin{aligned}\left\|\tilde{W}_i\right\| \leqslant & \eta_{i_1} e^{-\eta_{i_2} t} + \frac{\alpha_i}{m_{s_i} \eta_{i_2}} b_{\nabla\sigma_i} \left\|W_i\right\| \times \\ & \left(\left\|z_i(x_i, x_j, u_i, u_j, \omega_i, \omega_j)_{j \in N_i}\right\| \left\|\tilde{\varphi}_{\delta_i}\right\| + \left\|\varepsilon_{T_i}\right\|\right) \\ & + \frac{\alpha_i}{m_{s_i} \eta_{i_2}} \bar{e}_i,\ \ i = 1, \cdots, N \end{aligned} \quad (31)$$

for some $\eta_{i_1}, \eta_{i_2} > 0$, $i = 1, \cdots, N$ which is the desired result. This completes the proof. ■

5.2 *Actor Approximators for Optimal Control and Worst-Case Disturbance*

Based on (8) and (9), the estimates of control and worst-case disturbance policies can be approximated as follows,

$$\hat{u}_i = -(d_i + e_{i0}) {R_{ii}}^{-1} (\hat{\psi}_i \varsigma_i)^{\mathrm{T}} {\nabla\sigma_i}^{\mathrm{T}} \hat{W}_{i+N}, \quad (32)$$

$$\hat{\omega}_i = \frac{1}{\gamma^2} (d_i + e_{i0}) {T_{ii}}^{-1} (\hat{\beta}_i \vartheta_i)^{\mathrm{T}} {\nabla\sigma_i}^{\mathrm{T}} \hat{W}_{i+2N}, \quad (33)$$

where $\hat{W}_{i+N} \in \Re^{K_i}$, $\hat{W}_{i+2N} \in \Re^{K_i}$ denote the current estimated values of the ideal weight $W_i \in \Re^{K_i}$ by the actor approximators respectively. $\hat{\psi}_i$ and $\hat{\beta}_i$ are the estimated values of the ideal weights ${\psi_i}^{\star}$ and ${\beta_i}^{\star}$, $i = 1, \cdots, N$ respectively. Define the critic, and the actors weight estimation errors, $\tilde{W}_i, \tilde{W}_{i+N}, \tilde{W}_{i+2N}, \forall i$ respectively as,

$$\begin{aligned} \tilde{W}_i &= W_i - \hat{W}_i, \\ \tilde{W}_{i+N} &= W_i - \hat{W}_{i+N}, \\ \tilde{W}_{i+2N} &= W_i - \hat{W}_{i+2N}. \end{aligned} \quad (34)$$

In order to ensure closed-loop system stability and that the policies form a Nash equilibrium, the tuning laws for the two actors are selected as,

$$\begin{aligned} \dot{\hat{W}}_{i+N} = & -\alpha_{i+N}\{(S_i \hat{W}_{i+N} - F_i \hat{W}_i) \\ & -\hat{D}_i \hat{W}_{i+N} \frac{\bar{B}^{\mathrm{T}}_{i+N}}{2 m_{s_{i+N}}} \hat{W}_i \\ & - \sum_{j \in i_N} \hat{E}_j \hat{W}_{i+N} \frac{\bar{B}^{\mathrm{T}}_{j+N}}{2 m_{s_j}} \hat{W}_j\}, \end{aligned} \quad (35)$$

$$\begin{aligned} \dot{\hat{W}}_{i+2N} = & -\alpha_{i+2N}\{(\bar{S}_i \hat{W}_{i+N} - \bar{F}_i \hat{W}_i) \\ & + \frac{1}{\gamma^2} \hat{H}_i \hat{W}_{i+2N} \frac{\bar{B}^{\mathrm{T}}_{i+N}}{2 m_{s_{i+N}}} \hat{W}_i \\ & + \frac{1}{\gamma^2} \sum_{j \in i_N} \hat{G}_j \hat{W}_{i+2N} \frac{\bar{B}^{\mathrm{T}}_{j+N}}{2 m_{s_j}} \hat{W}_j\} \end{aligned} \quad (36)$$

where

$$\begin{aligned} & \alpha_{i+N} > 0,\ \alpha_{i+2N} > 0, \\ & \hat{E}_j = {c_j}^2 \nabla\sigma_j (\hat{\psi}_j \varsigma_j) {R_{jj}}^{-T} R_{ij} {R_{jj}}^{-1} (\hat{\psi}_j \varsigma_j)^{\mathrm{T}} {\nabla\sigma_j}^{\mathrm{T}}, \\ & \hat{D}_i = {c_i}^2 \nabla\sigma_i (\hat{\psi}_i \varsigma_i) {R_{ii}}^{-T} (\hat{\psi}_i \varsigma_i)^{\mathrm{T}} {\nabla\sigma_i}^{\mathrm{T}}, \\ & \hat{H}_i = {c_i}^2 \nabla\sigma_i (\hat{\beta}_i \vartheta_i) {T_{ii}}^{-T} (\hat{\beta}_i \vartheta_i)^{\mathrm{T}} {\nabla\sigma_i}^{\mathrm{T}}, \\ & \hat{G}_j = {c_j}^2 \nabla\sigma_j (\hat{\beta}_j \vartheta_j) {T_{jj}}^{-T} T_{ij} {T_{jj}}^{-1} (\hat{\beta}_j \vartheta_j)^{\mathrm{T}} {\nabla\sigma_j}^{\mathrm{T}}, \\ & \hat{B}_{i+N} = \nabla\sigma_i (\hat{\varphi}_i \left[z_i(x_i, x_j, \hat{u}_i, \hat{u}_{N_i}, \hat{\omega}_i, \hat{\omega}_{N_i})\right]), \\ & m_{s_{i+N}} = 1 + \hat{B}^{\mathrm{T}}_{i+N} \hat{B}_{i+N}, i_{N_i} = \{j : (i,j) \in \Sigma\}, \\ & \bar{B}_{i+N} = \frac{\hat{B}_{i+N}}{1 + \hat{B}^{\mathrm{T}}_{i+N} \hat{B}_{i+N}},\ \ c_i = (d_i + e_{i0}),\ i = 1, \cdots, N. \end{aligned}$$

$S_i \in \Re^{K_i \times K_i}$, $F_i \in \Re^{K_i \times K_i}$, $\bar{S}_i \in \Re^{K_i \times K_i}$, $\bar{F}_i \in \Re^{K_i \times K_i}$ are diagonal positive definite tuning matrices.

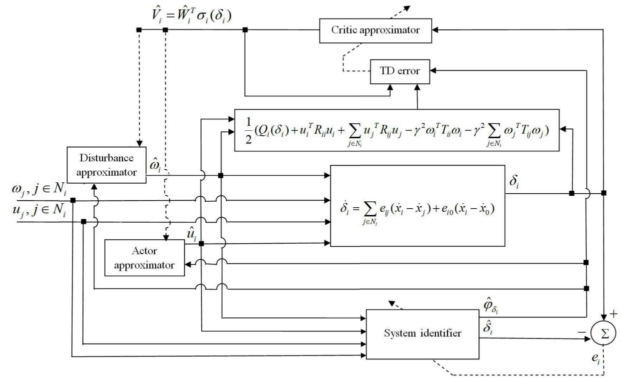

Fig. 1: Optimal distributed disturbance rejection algorithm for every agent i, under unknown dynamics.

Finally the proposed method can be summarized in the following algorithm.

Algorithm 1: Disturbance Rejection Algorithm in Nonlinear Networked Games with Unknown Dynamics

---

1) Initialize control ${u_i}^0$ by the initial actor weights $\hat{W}^0_{i+N}$, disturbance ${\omega_i}^0$ by the initial disturbance weights $\hat{W}^0_{i+2N}$ and value function V^0_i by the initial critic weights $\hat{W}^0_i$ and initialize the parameters of unknown dynamics $\hat{\varphi}^0_{\delta_i}, \forall i = 1, \cdots, N$.

For $\forall k = 0, 1, \cdots$

2) Update the identified parameters of the unknown dynamics $\hat{\varphi}^k_{\delta_i}$ using (17).

3) Update the critic value function approximation weights $\hat{W}^k_i$ through the gradient-based tuning law (25).

4) Update the control $\hat{u}^k_i = -(d_i + e_{i0}) {R_{ii}}^{-1} (\hat{\psi}^k_i \varsigma^k_i)^{\mathrm{T}} {\nabla\sigma_i}^{\mathrm{T}} \hat{W}^k_{i+N}$ and disturbance $\hat{\omega}^k_i = \frac{1}{\gamma^2} (d_i + e_{i0}) {T_{ii}}^{-1} (\hat{\beta}^k_i \vartheta^k_i)^{\mathrm{T}} {\nabla\sigma_i}^{\mathrm{T}} \hat{W}^k_{i+2N}$ policies through the actors weights $\hat{W}^k_{i+N}$ and $\hat{W}^k_{i+2N}$ tuning laws from (35) and (36) respectively.

5) Go to step 3 until convergence.

End for

The block diagram of the proposed distributed learning algorithm for every agent is depicted in Fig. 1 where the solid lines show the associated signals and the dashed lines shows approximators weights tunings.

5.3 *Stability and Convergence Analysis*

The main theorem which provides, closed-loop system stability and convergence of the policies to a Nash equilibrium is now presented.

Theorem 2. *Consider the dynamical system (19) with ${\theta_i}^{\star}$, $\psi^{\star}_i$, $\psi^{\star}_j|_{j \in N_i}$, $\beta^{\star}_i$, and $\beta^{\star}_j|_{j \in N_i}, i = 1, \cdots, N$ unknown. Assume that $\bar{B}_{i+N}$ is PE and that Assumptions 1-3 hold. Let the approximator identifiers weights be updated by (17), the value function, control and worst-case disturbance of each agent be respectively given by (23), (32) and (33) and that the tuning laws of agent i critic, the optimal control actor and the worst-case disturbance actor, are respectively given by (25), (35) and (36). Then the closed-loop system states $\delta_i(t)$, the critic approximators errors $\tilde{W}_i$, the actor approximator errors $\tilde{W}_{i+N}$ and the disturbance approximator errors $\tilde{W}_{i+2N}$ are UUB, for a sufficiently large number of approximators basis.*

Proof: See the Appendix. □

IET Research Journals, pp. 1–11

Corollary 1. *Let the assumptions and statements of Theorem 2 hold. Then the solution provided by the critic, and the two actor approximators provides a solution to the coupled HJI equations (10). The inputs $\hat{u}_i, \hat{\omega}_i, i = 1, \cdots, N$ form a Nash equilibrium solution of the zero-sum game.*

Proof: The proof is a direct consequence of Theorem 2. ∎

6 Simulations

Consider a network of 3 single-link manipulators with revolute joints actuated by a DC motor (as shown in Figure 2). Every single-link manipulator states are motor position and velocity, and the link position and velocity [28,29].The pinning gain and the edge weights are chosen as 1. In the graph structure of Figure 2, the manipulators' dynamics for $i = 1, 2, 3$ are as follows,

$$\dot{x}_i = \begin{pmatrix} p_1 x_{i2} \\ p_2 x_{i1} + p_3 x_{i2} + p_4 x_{i3} \\ p_5 x_{i4} \\ p_6 x_{i1} + p_7 x_{i3} + p_8 \sin x_{i3} \end{pmatrix} + \begin{bmatrix} 0 \\ p_9 u_i \\ p_{10}\omega_i \\ 0 \end{bmatrix},$$

$$x_i = \begin{bmatrix} x_{i1} \\ x_{i2} \\ x_{i3} \\ x_{i4} \end{bmatrix}, \dot{x}_0 = \begin{pmatrix} p_1 x_{02} \\ p_2 x_{01} + p_3 x_{02} + p_4 x_{03} \\ p_5 x_{04} \\ p_6 x_{01} + p_7 x_{03} + p_8 \sin x_{03} \end{pmatrix}.$$

It is assumed that all the agents structures are known and that the MAS parameters vector $P = [p_1, p_2, \cdots, p_{10}]$ is unknown. The unknown identifier weights are considered to be $P = [1, -48.6, -1.25, 48.6, 1, 19.5, -19.5, -3.33, 21.6, 1]$. As pointed out in Remark 5, in order to reduce the system identification error bound in (18), we pick $a = 40,\ p_i = 1000,\ \Gamma_i = 100, i = 1, 2, 3$. The sample rate constant is chosen as 0.001 second in order to satisfy the rank condition and limit the identification error bound (18) for every agent. We pick $Q_i(\delta_i) = {\delta_i}^{\mathrm{T}} Q_{ii} \delta_i,\ \ Q_{ii} = I, R_{ii} = 10,\ R_{ij} = 1, T_{ii} = 10,\ T_{ij} = 1,\ (i \neq j,\ j \in N_i)$ and $\gamma = 5$, for $i = 1, 2, 3$. In order to guarantee that inequality (43) holds the design parameters are selected as $S_i = F_i = \bar{S}_i = \bar{F}_i = 100I$. Since the critic approximators are needed to be faster than the actors to guarantee closed-loop stability, we pick $\alpha_i = 10, \alpha_{i+N} = \alpha_{i+2N} = 1$. The critic and the actor approximators activation functions are chosen as $\sigma_i = [{\delta_{i1}}^2, \delta_{i1}\delta_{i2}, \delta_{i1}\delta_{i3}, \delta_{i1}\delta_{i4}, {\delta_{i2}}^2, \delta_{i2}\delta_{i3}, \delta_{i2}\delta_{i4}, {\delta_{i3}}^2, \delta_{i3}\delta_{i4}, {\delta_{i4}}^2], i = 1, 2, 3$. The PE is guaranteed by adding a small exponentially decreasing probing noise to the control inputs. Figure 3 shows that the unknown parameters of the agents' dynamics converge to their true values. The evolution of the critic weights is shown in Figure 4. Figure 5 shows the local tracking errors and their convergence to a neighborhood of zero.

Since the motor and the link positions and velocities are limited, the state vector $x_i(t)$ is limited to a compact set Ω_i $(x_i \in \Omega_i)$ and hence it can be inferred that Assumptions 1-3 are satisfied. As shown in Figure 3, the unknown parameters uniformly converge to their true values and considering the chosen σ_i and consequent $\nabla\sigma_i$, it is evident that Assumption 2 is satisfied. Also the obtained $\hat{V}_i = \hat{W}_i^{\mathrm{T}} \sigma_i(\delta_i)$ is a local smooth solution which indicates the satisfaction of Assumption 1. Assumption 3 is also satisfied since in every compact set $\Omega_i, \forall i, \|f_i(x_i)\|, \|g_i(x_i)\|$ and $\|k_i(x_i)\|$ are all bounded and as it is depicted in Figure 5, there is a constant matrix $W_{i\ \max}$ so that $\|W_i\| < W_{i\ \max}$.

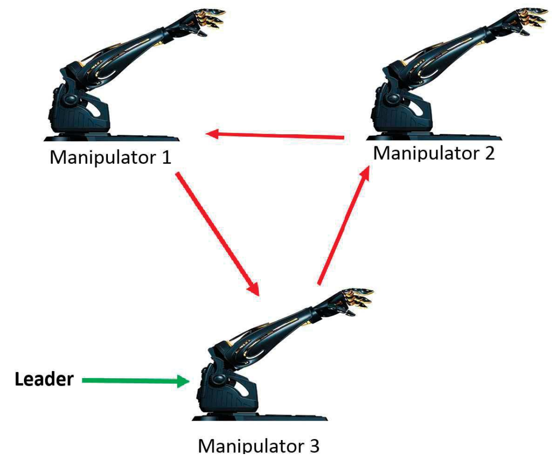

Fig. 2: A MAS of 3 single-link manipulators.

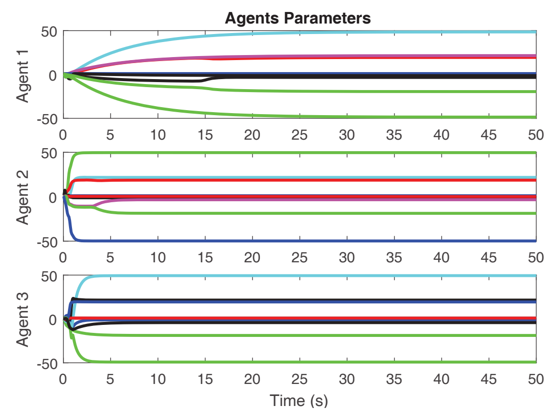

Fig. 3: The evolution of the unknown parameters of the manipulators.

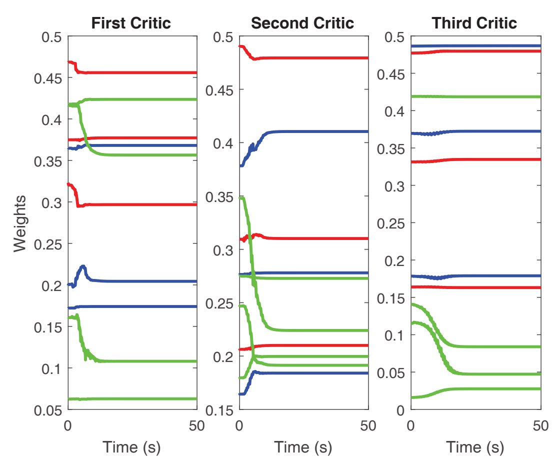

Fig. 4: The evolution and convergence of the critic weights.

7 Conclusion

An online distributed learning control algorithm based on RL techniques is presented to solve the continuous-time unknown multi-player nonlinear graphical games in the presence of disturbances. The distributed learning algorithm is implemented in the form of actor-critic-disturbance structures to approximate the optimal policies of the players. In order to identify the unknown dynamics, we use identifiers in conjunction with the actor-critic-disturbance networks. The coupled HJI equations of the agents are approximately solved. The boundedness of the closed-loop signals are proved according to Lyapunov stability and it is ensured that the policies form a Nash equilibrium. Future research efforts will focus on extending the model-based technique to a model-free approach.

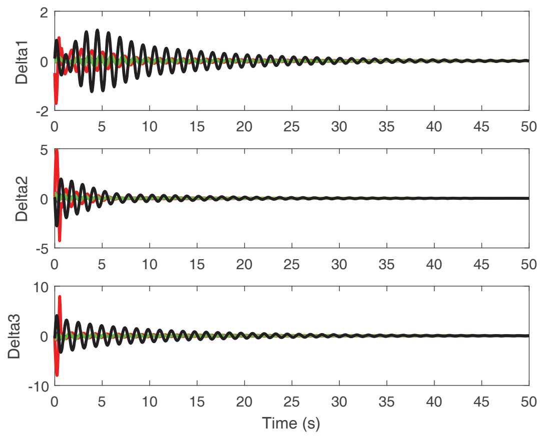

Fig. 5: The evolution of the manipulators' tracking errors.

8 Acknowledgments

This work is supported in part by the National Science Foundation under grant NSF CAREER CPS-1851588, in part by NATO under grant No. SPS G5176, in part by ONR Minerva under grant No. $N00014-18-1-2160$.

9 References

[1] Dimarogonas, D. V., Frazzoli, E., Johansson, K. H.: 'Distributed Event-Triggered Control for Multi-Agent Systems', IEEE Transactions on Automatic Control, 2012, 57, (5), pp. 1291–1297.

[2] Li, Z., Ren, W., Liu, X., Fu, M.: 'Consensus of Multi-Agent Systems with General Linear and Lipschitz Nonlinear Dynamics Using Distributed Adaptive Protocols', IEEE Transactions on Automatic Control, 2013, 58, (7), pp. 1786–1791.

[3] Yang, T., Roy, S., Wan, Y., et al.: 'Constructing consensus controllers for networks with identical general linear agents', International Journal of Robust and Nonlinear Control, 2011, 21, pp: 1237–1256

[4] Meng, Z., Ren, W., Cao, Y., et al.: 'Leaderless and Leader-Following Consensus With Communication and Input Delays Under a Directed Network Topolog', IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics, 2011, 41, (1), pp. 75 – 88

[5] Rezaei, M. H., Menhaj, M. B.: 'Adaptive output stationary average consensus for heterogeneous unknown linear multi-agent systems', IET Control Theory and Applications, 2018, 12, (7), pp. 847–856

[6] Li, C. J., Liu, G. P. :'Data-driven consensus for non-linear networked multi-agent systems with switching topology and time-varying delays', IET Control Theory and Applications, 2018, 12, (12), pp. 1773–1779

[7] Vamvoudakis, K. G., Lewis, F. L., Hudas, G.R.: 'Multi-agent differential graphical games: Online adaptive learning solution for synchronization with optimality', Automatica, 2012, 48, pp.1598–1611

[8] Tatari, F., Naghibi-Sistani, M. B., Vamvoudakis, K. G.: 'Distributed Learning Algorithm for Nonlinear Differential Graphical Games', Transactions of the Institute of Measurement and Control, 2017, 39, (2), pp. 173–182

[9] Jiao, Q., Modares, H.,Xu, S., et al.: 'Multi-Agent zero-sum differential graphical games for disturbance rejection in distributed control', Automatica, 2016, 69, pp. 24–34

[10] Kamalapurkar, R., Klotz, J. R., Walters, P., et al.: 'Model-based reinforcement learning in differential graphical games', IEEE Transactions on Control of Network Systems, 2018, 5, (1), pp. 423–433

[11] Tatari, F., Naghibi-Sistani, M. B., Vamvoudakis, K. G.: 'Distributed Optimal Synchronization Control of Linear Networked Systems under Unknown Dynamics'. Proc. American Control Conference, 2017, Seattle, WA, pp. 668–673

[12] Mazouchi, M., Naghibi-S., M. B., Hosseini-S., S. K.: 'A novel distributed optimal adaptive control algorithm for nonlinear multi-agent differential graphical games', 2018, IEEE/CAA Journal of Automatica Sinica, 5, (1), pp. 331–341.

[13] Sutton, R.S., Barto, A.G.: 'Reinforcement learning an introduction' (MIT Press,1998, 1nd edn.)

[14] Gao, W., Jiang, Z.: 'Adaptive dynamic programming and adaptive optimal output regulation of linear systems', IEEE Transactions on Automatic Control, 2016, 61, (12), pp. 4164–4169.

[15] Modares, H., Lewis, F. L., Naghibi-Sistani, M. B.: 'Online solution of nonquadratic two-player zero-sum games arising in the HâĹđ control of constrained-input systems', International Journal of Adaptive Control and Signal Processing, 2014, 28, pp. 232–254.

[16] Vamvoudakis, K. G., Lewis, F. L.: 'Online solution of nonlinear two-player zero-sum games using synchronous policy iteration', International Journal of Robust and Nonlinear Control, 2012, 22, (13), pp. 1460–1483.

[17] Luy, N. T., Thanh, N.T., Tri, H. M.: 'Reinforcement learning-based robust adaptive tracking control for multi-wheeled mobile robots synchronization with optimality'. In IEEE workshop on robotic intelligence in informationally structured space, 2013, pp. 74–81

[18] Tan, L. N.: 'Distributed optimal integrated tracking control for separate kinematic and dynamic uncertain non-holonomic mobile mechanical multi-agent systems', 2017, IET Control Theory and Applications, 11, (18), pp. 3249–3260.

[19] Aliyu, M.D.S.: 'Nonlinear H_∞ control, Hamiltonian systems and Hamilton-Jacobi equations' (CRC Press, Taylor and Francis Group, Boca Raton, 2011)

[20] Basar, T., Bernhard, P.: ' H_∞ optimal control and related minimax design problems: A dynamic game approach' (Springer, 2008, 2nd edn. 2008)

[21] Lewis, F. L., Jagannathan, S., Yesildirek, A.: 'Neural Network Control of Robot Manipulators and Nonlinear Systems', London, UK: Taylor and Francis, 1999.

[22] Abu-Khalaf, M., Lewis, F. L.: 'Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach', Automatica, 2005, 41, pp. 779–791

[23] Modares, H., Lewis, F. L., Naghibi-Sistani, M.B.: 'Adaptive Optimal Control of Unknown Constrained-Input Systems Using Policy Iteration and Neural Networks', IEEE Transactions on Neural Networks and Learning Systems, 2013, 24, (10), pp. 1513–1525

[24] Vamvoudakis,K. G.: 'Non-Zero Sum Nash Q-learning for Unknown Deterministic Continuous-Time Linear Systems', Automatica, 2015, 61, pp. 274–281.

[25] Slotine, J. J. E., Li, W. P.: 'Applied Nonlinear Control', Englewood Cliffs, NJ, USA: Prentice Hall, 1991.

[26] Sastry, S., Bodson, M.: 'Adaptive Control: Stability, Convergence, and Robustness', Englewood Cliffs, NJ: Prentice Hall, 1989.

[27] Ioannou, P., Sun, J.: 'Robust Adaptive Control', (1996, Prentice Hall, New Jersey.)

[28] Spong, M.: 'Modeling and control of elastic joint robots', ASME Journal of Dynamic Systems, Measurement and Control, 1987, 109, pp. 31–319.

[29] Raghavan, S.: 'Observers and compensators for nonlinear systems, with application to exible-joint robots', Ph.D. Dissertation, University of California at Berkeley, Berkeley, 1992, CA 94720, U.S.A.

[30] Hardy, G., Littlewood, J., Polya, G.: 'Inequalities', (Cambridge Univ. Press, 1989, 2nd edn. 1989)

[31] Khalil, H.K.: ' Nonlinear systems', (Prentice-Hall, 1996)

10 Appendix

Proof of Theorem 2

The following fact will be used in the proof.

Fact 1. *For any two vectors x and y and any $\varepsilon > 0$, it holds that* $x^{\mathrm{T}}y \leqslant \varepsilon\frac{\|x\|^2}{2} + \frac{\|y\|^2}{2\varepsilon}$. □

We consider the following Lyapunov function,

$$L(t) = \sum_{i=1}^{N}\{V_i(t) + \overbrace{\frac{1}{2}\tilde{W}_i^{\mathrm{T}}{\alpha_i}^{-1}\tilde{W}_i}^{L_i(t)} + \overbrace{\frac{1}{2}\tilde{W}_{i+N}^{\mathrm{T}}{\alpha_{i+N}}^{-1}\tilde{W}_{i+N}}^{L_{i+N}(t)} + \overbrace{\frac{1}{2}\tilde{W}_{i+2N}^{\mathrm{T}}{\alpha_{i+2N}}^{-1}\tilde{W}_{i+2N}}^{L_{i+2N}(t)}\}. \tag{37}$$

The time derivative of Lyapunov function is given by

$$\dot{L}(t) = \sum_{i=1}^{N}\{\dot{V}_i(t) \overbrace{-\tilde{W}_i^{\mathrm{T}}{\alpha_i}^{-1}\dot{\hat{W}}_i}^{\dot{L}_i(t)} \overbrace{-\tilde{W}_{i+N}^{\mathrm{T}}{\alpha_{i+N}}^{-1}\dot{\hat{W}}_{i+N}}^{\dot{L}_{i+N}(t)} \overbrace{-\tilde{W}_{i+2N}^{\mathrm{T}}{\alpha_{i+2N}}^{-1}\dot{\hat{W}}_{i+2N}}^{\dot{L}_{i+2N}(t)}\}. \tag{38}$$

The first term in (38), using (16), (20), (32)-(34) and doing some simple algebra, can be written as

$$\begin{aligned}\sum_{i=1}^{N}\dot{V}_i(t) = \sum_{i=1}^{N}\{\nabla V_i(\dot{\delta}_i(t))\} = \sum_{i=1}^{N}\{\dot{\bar{L}}_{V_i} \\ -\frac{1}{2}Q_i(\delta_i) - {W_i}^{\mathrm{T}}\nabla\sigma_i\{{c_i}^2({\psi_i}^{\star}\varsigma_i + \varepsilon_{g_i})\times \\ {R_{ii}}^{-1}((\tilde{\psi}_i + {\psi_i}^{\star})\varsigma_i)^{\mathrm{T}}\nabla{\sigma_i}^{\mathrm{T}}W_i \\ -{c_i}^2({\psi_i}^{\star}\varsigma_i + \varepsilon_{g_i}){R_{ii}}^{-1}((\tilde{\psi}_i + {\psi_i}^{\star})\varsigma_i)^{\mathrm{T}}\times \\ \nabla{\sigma_i}^{\mathrm{T}}\tilde{W}_{i+N} + \sum_{j\in N_i}e_{ij}(c_j({\psi_j}^{\star}\varsigma_j + \varepsilon_{g_j}){R_{jj}}^{-1}\times \\ ((\tilde{\psi}_j + {\psi_j}^{\star})\varsigma_j)^{\mathrm{T}}\nabla{\sigma_j}^{\mathrm{T}}W_j \\ -c_j({\psi_j}^{\star}\varsigma_j + \varepsilon_{g_j}){R_{jj}}^{-1}((\tilde{\psi}_j + {\psi_j}^{\star})\varsigma_j)^{\mathrm{T}}\times \\ \nabla{\sigma_j}^{\mathrm{T}}\tilde{W}_{j+N}) - \frac{1}{\gamma^2}{c_i}^2({\beta_i}^{\star}\vartheta_i + \varepsilon_{k_i}){T_{ii}}^{-1}\times \\ ((\tilde{\beta}_i + {\beta_i}^{\star})\vartheta_i)^{\mathrm{T}}\nabla{\sigma_i}^{\mathrm{T}}W_i + \frac{1}{\gamma^2}{c_i}^2\times \\ ({\beta_i}^{\star}\vartheta_i + \varepsilon_{k_i}){T_{ii}}^{-1}((\tilde{\beta}_i + {\beta_i}^{\star})\vartheta_i)^{\mathrm{T}}\nabla{\sigma_i}^{\mathrm{T}}\times \\ \tilde{W}_{i+2N} - \frac{1}{\gamma^2}\sum_{j\in N_i}e_{ij}(c_j({\beta_j}^{\star}\vartheta_j + \varepsilon_{k_j}){T_{jj}}^{-1}\times \\ ((\tilde{\beta}_j + {\beta_j}^{\star})\vartheta_j)^{\mathrm{T}}\nabla{\sigma_j}^{\mathrm{T}}W_j + \frac{1}{\gamma^2}c_j\times \\ ({\beta_j}^{\star}\vartheta_j + \varepsilon_{k_j}){T_{jj}}^{-1}((\tilde{\beta}_j + {\beta_j}^{\star})\vartheta_j)^{\mathrm{T}}\times \\ \nabla{\sigma_j}^{\mathrm{T}}\tilde{W}_{j+2N})\}\} + \sum_{i=1}^{N}v_{i0},\end{aligned} \tag{39}$$

where,

$$\begin{aligned}\dot{\bar{L}}_{V_i} = -{W_i}^{\mathrm{T}}\frac{\partial\sigma_i}{\partial\delta_i}[-(d_i + e_{i0})^2({\psi_i}^{\star}\varsigma_i + \varepsilon_{g_i}){R_{ii}}^{-1}\times \\ ({\psi_i}^{\star}\varsigma_i + \varepsilon_{g_i})^{\mathrm{T}}\nabla{\sigma_i}^{\mathrm{T}}W_i + \sum_{j\in N_i}e_{ij}(d_j + e_{j0})\times \\ ({\psi_j}^{\star}\varsigma_j + \varepsilon_{g_j}){R_{ii}}^{-1}({\psi_j}^{\star}\varsigma_j + \varepsilon_{g_j})^{\mathrm{T}}\nabla{\sigma_j}^{\mathrm{T}}W_j] \\ -\frac{1}{2}(d_i + e_{i0})^2{W_i}^{\mathrm{T}}\nabla\sigma_i({\psi_i}^{\star}\varsigma_i + \varepsilon_{g_i}){R_{ii}}^{-T}\times \\ ({\psi_i}^{\star}\varsigma_i + \varepsilon_{g_i})^{\mathrm{T}}\nabla{\sigma_i}^{\mathrm{T}}W_i - \frac{1}{2}\sum_{j\in N_i}(d_j + e_{j0})^2\times \\ {W_j}^{\mathrm{T}}\nabla\sigma_j({\psi_j}^{\star}\varsigma_j + \varepsilon_{g_j}){R_{jj}}^{-T}R_{ij}{R_{jj}}^{-1}\times \\ ({\psi_j}^{\star}\varsigma_j + \varepsilon_{g_j})^{\mathrm{T}}\nabla{\sigma_j}^{\mathrm{T}}W_j + e_{B_i},\end{aligned}$$

$$\begin{aligned}\sum_{i=1}^{N}v_{i0} = \sum_{i=1}^{N}\{\nabla v_i[\sum_{j\in N_i}e_{ij}(f_i(x_i) - f_j(x_j)) + e_{i0}(f_i(x_i) \\ - f_0(x_0)) - {c_i}^2g_i(x_i){R_{ii}}^{-1}(\hat{\psi}_i\varsigma_i)^{\mathrm{T}}\nabla{\sigma_i}^{\mathrm{T}}\hat{W}_{i+N} \\ + \sum_{j\in N_i}e_{ij}c_jg_j(x_j){R_{jj}}^{-1}(\hat{\psi}_j\varsigma_j)^{\mathrm{T}}\nabla{\sigma_j}^{\mathrm{T}}\hat{W}_{j+N} \\ + \frac{1}{\gamma^2}{c_i}^2k_i(x_i){T_{ii}}^{-1}(\hat{\beta}_i\vartheta_i)^{\mathrm{T}}\nabla{\sigma_i}^{\mathrm{T}}\hat{W}_{i+2N} - \frac{1}{\gamma^2}\times \\ \sum_{j\in N_i}e_{ij}c_jk_j(x_j){T_{jj}}^{-1}(\hat{\beta}_j\vartheta_j)^{\mathrm{T}}\nabla{\sigma_j}^{\mathrm{T}}\hat{W}_{j+2N}]\}.\end{aligned}$$

For the second term in (38), $\dot{L}_i(t)$, from (20), (25) and (34), we have

$$\begin{aligned}
\dot{L}_i = & -\tilde{W}_i^{\mathrm{T}}\alpha_i{}^{-1}\dot{\hat{W}}_i = \tilde{W}_i^{\mathrm{T}}\frac{\hat{B}_{i+N}}{(1+\hat{B}_{i+N}^{\mathrm{T}}\hat{B}_{i+N})^2}\times \\
& \{W_i{}^{\mathrm{T}}\nabla\sigma_i(\tilde{\varphi}_i z_i(x_i,x_j,\hat{u}_i,\hat{u}_{N_i})) \\
& -\hat{B}_{i+N}^{\mathrm{T}}\tilde{W}_i + \frac{1}{2}(W_i-\tilde{W}_{i+N})^{\mathrm{T}}\hat{D}_i(W_i-\tilde{W}_{i+N})\times \\
& +\frac{1}{2}\sum_{j\in N_i}(W_j-\tilde{W}_{j+N})^{\mathrm{T}}\hat{E}_j(W_j-\tilde{W}_{j+N}) - \frac{1}{2\gamma^2}\times \\
& (W_i-\tilde{W}_{i+2N})^{\mathrm{T}}\hat{H}_i(W_i-\tilde{W}_{i+2N}) - \frac{1}{2\gamma^2}\times \\
& \sum_{j\in N_i}(W_j-\tilde{W}_{j+2N})^{\mathrm{T}}\hat{G}_j(W_j-\tilde{W}_{j+2N}) - W_i{}^{\mathrm{T}}\nabla\sigma_i\times \\
& (c_i(\psi_i{}^{\star}\varsigma_i)(u_i-\hat{u}_i) - \sum_{j\in N_i}e_{ij}(\psi_j{}^{\star}\varsigma_j)(u_j-\hat{u}_j) \\
& -\frac{1}{\gamma^2}c_i(\beta_i{}^{\star}\vartheta_i)(\omega_i-\hat{\omega}_i) + \frac{1}{\gamma^2}\sum_{j\in N_i}e_{ij}(\beta_j{}^{\star}\vartheta_j)(\omega_j-\hat{\omega}_j)) \\
& -W_i{}^{\mathrm{T}}\nabla\sigma_i\varepsilon_{T_i} + \varepsilon_{L_i} + e_{B_i}\}, \qquad (40)
\end{aligned}$$

where

$$\begin{aligned}
\varepsilon_{L_i} = & -\frac{1}{2}(d_i+e_{i0})^2 W_i{}^{\mathrm{T}}\nabla\sigma_i(\psi_i{}^{\star}\varsigma_i+\varepsilon_{g_i})\times \\
& R_{ii}{}^{-T}(\psi_i{}^{\star}\varsigma_i+\varepsilon_{g_i})^{\mathrm{T}}\nabla\sigma_i{}^{\mathrm{T}}W_i \\
& -\frac{1}{2}\sum_{j\in N_i}(d_j+e_{j0})^2 W_j{}^{\mathrm{T}}\nabla\sigma_j(\psi_j{}^{\star}\varsigma_j+\varepsilon_{g_j})\times \\
& R_{jj}{}^{-T}R_{ij}R_{jj}{}^{-1}(\psi_j{}^{\star}\varsigma_j+\varepsilon_{g_j})^{\mathrm{T}}\nabla\sigma_j{}^{\mathrm{T}}W_j.
\end{aligned}$$

Using (25), (34) and Fact 1, (40) is written as,

$$\begin{aligned}
\dot{L}_i(t) \leqslant & \dot{L}_i + \tilde{W}_{i+N}^{\mathrm{T}}\hat{D}_i\tilde{W}_{i+N}\frac{\bar{B}_{i+N}^{\mathrm{T}}}{2m_{s_{i+N}}}\tilde{W}_i + \sum_{j\in N_i}\tilde{W}_{j+N}^{\mathrm{T}}\times \\
& \hat{E}_j\tilde{W}_{j+N}\frac{\bar{B}_{i+N}^{\mathrm{T}}}{2m_{s_{i+N}}}\tilde{W}_i - \frac{1}{\gamma^2}\tilde{W}_{i+2N}^{\mathrm{T}}\hat{H}_i\tilde{W}_{i+2N}\frac{\bar{B}_{i+N}^{\mathrm{T}}}{2m_{s_{i+N}}}\tilde{W}_i \\
& -\frac{1}{\gamma^2}\sum_{j\in N_i}\tilde{W}_{j+2N}^{\mathrm{T}}\hat{G}_j\tilde{W}_{j+2N}\frac{\bar{B}_{i+N}^{\mathrm{T}}}{2m_{s_{i+N}}}\tilde{W}_i, \qquad (41)
\end{aligned}$$

where

$$\begin{aligned}
\dot{L}_i = & -r_i\tilde{W}_i^{\mathrm{T}}\bar{B}_{i+N}\bar{B}_{i+N}^{\mathrm{T}}\tilde{W}_i + \left\|\tilde{W}_i^{\mathrm{T}}\frac{\bar{B}_{i+N}}{m_{s_{i+N}}}\right\|(a_{i_1}+k_{T_i}) \\
& +\frac{\varepsilon}{2}a_{i_2}\delta_i{}^{\mathrm{T}}\delta_i - \frac{k_{N_i}}{2}W_i\tilde{W}_{i+N}\frac{\bar{B}_{i+N}^{\mathrm{T}}}{m_{s_{i+N}}}\tilde{W}_i - \frac{k_{N_i}}{4}\tilde{W}_{i+N}\times \\
& W_i\frac{\bar{B}_{i+N}^{\mathrm{T}}}{m_{s_{i+N}}}\tilde{W}_i - \sum_{j\in N_i}\frac{k_{O_j}}{4}\times W_j{}^{\mathrm{T}}\tilde{W}_{j+N}\frac{\bar{B}_{i+N}^{\mathrm{T}}}{m_{s_{i+N}}}\tilde{W}_i \\
& -\sum_{j\in N_i}\frac{k_{O_j}}{4}\tilde{W}_{j+N}^{\mathrm{T}}W_j\frac{\bar{B}_{i+N}^{\mathrm{T}}}{m_{s_{i+N}}}\tilde{W}_i + \frac{k_{P_i}}{2\gamma^2}W_i\tilde{W}_{i+2N}\times \\
& \frac{\bar{B}_{i+N}^{\mathrm{T}}}{m_{s_{i+N}}}\tilde{W}_i + \frac{k_{P_i}}{4\gamma^2}\tilde{W}_{i+2N}W_i\frac{\bar{B}_{i+N}^{\mathrm{T}}}{m_{s_{i+N}}}\tilde{W}_i + \sum_{j\in N_i}\frac{k_{Q_j}}{4\gamma^2}\times \\
& W_j{}^{\mathrm{T}}\tilde{W}_{j+2N}\frac{\bar{B}_{i+N}^{\mathrm{T}}}{m_{s_{i+N}}}\tilde{W}_i + \sum_{j\in N_i}\frac{k_{Q_j}}{4\gamma^2}\tilde{W}_{j+2N}^{\mathrm{T}}W_j\times \\
& \frac{\bar{B}_{i+N}^{\mathrm{T}}}{m_{s_{i+N}}}\tilde{W}_i + c_i{}^2W_i{}^{\mathrm{T}}\nabla\sigma_i(\tilde{\psi}_i\varsigma_i)R_{ii}^{-1}(\psi_i^{\star}\varsigma_i)^{\mathrm{T}}\nabla\sigma_i{}^{\mathrm{T}}\times \\
& (\tilde{W}_{i+N}-W_i)\frac{\bar{B}_{i+N}^{\mathrm{T}}}{m_{s_{i+N}}}\tilde{W}_i - W_i^{\mathrm{T}}\tilde{D}_iW_i\frac{\bar{B}_{i+N}^{\mathrm{T}}}{m_{s_{i+N}}}\tilde{W}_i \\
& +W_i^{\mathrm{T}}\tilde{D}_i\tilde{W}_{i+N}\frac{\bar{B}_{i+N}^{\mathrm{T}}}{m_{s_{i+N}}}\tilde{W}_i + W_i{}^{\mathrm{T}}\nabla\sigma_i\sum_{j\in N_i}c_je_{ij}\times \\
& (\tilde{\psi}_j\varsigma_j)R_{jj}^{-1}(\psi_j^{\star}\varsigma_j)^{\mathrm{T}}\nabla\sigma_j{}^{\mathrm{T}}(W_j-\tilde{W}_{j+N})\frac{\bar{B}_{i+N}^{\mathrm{T}}}{m_{s_{i+N}}}\tilde{W}_i \\
& +W_i{}^{\mathrm{T}}\nabla\sigma_i\sum_{j\in N_i}c_je_{ij}(\tilde{\psi}_j\varsigma_j)R_{jj}^{-1}(\tilde{\psi}_j\varsigma_j)^{\mathrm{T}}\nabla\sigma_j{}^{\mathrm{T}}\times \\
& (W_j-\tilde{W}_{j+N})\frac{\bar{B}_{i+N}^{\mathrm{T}}}{m_{s_{i+N}}}\tilde{W}_i - \frac{1}{\gamma^2}c_i{}^2W_i{}^{\mathrm{T}}\nabla\sigma_i(\tilde{\beta}_i\vartheta_i)T_{ii}^{-1}\times \\
& (\beta_i^{\star}\vartheta_i)^{\mathrm{T}}\nabla\sigma_i{}^{\mathrm{T}}(\tilde{W}_{i+2N}-W_i)\frac{\bar{B}_{i+N}^{\mathrm{T}}}{m_{s_{i+N}}}\tilde{W}_i + \frac{1}{\gamma^2}W_i^{\mathrm{T}}\tilde{H}_i\times \\
& W_i\frac{\bar{B}_{i+N}^{\mathrm{T}}}{m_{s_{i+N}}}\tilde{W}_i - \frac{1}{\gamma^2}W_i^{\mathrm{T}}\tilde{H}_i\tilde{W}_{i+2N}\frac{\bar{B}_{i+N}^{\mathrm{T}}}{m_{s_{i+N}}}\tilde{W}_i \\
& -\frac{1}{\gamma^2}W_i{}^{\mathrm{T}}\nabla\sigma_i\sum_{j\in N_i}c_je_{ij}(\tilde{\beta}_j\vartheta_j)T_{jj}^{-1}(\beta_j^{\star}\vartheta_j)^{\mathrm{T}}\nabla\sigma_j{}^{\mathrm{T}}\times \\
& (W_j-\tilde{W}_{j+2N})\frac{\bar{B}_{i+N}^{\mathrm{T}}}{m_{s_{i+N}}}\tilde{W}_i \\
& -\frac{1}{\gamma^2}W_i{}^{\mathrm{T}}\nabla\sigma_i\sum_{j\in N_i}c_je_{ij}(\tilde{\beta}_j\vartheta_j)T_{jj}^{-1} \\
& -\frac{1}{\gamma^2}W_i{}^{\mathrm{T}}\nabla\sigma_i(\tilde{\beta}_j\vartheta_j)^{\mathrm{T}}\nabla\sigma_j{}^{\mathrm{T}}(W_j-\tilde{W}_{j+2N})\frac{\bar{B}_{i+N}^{\mathrm{T}}}{m_{s_{i+N}}}\tilde{W}_i \\
& +(e_{B_i}+\varepsilon_{L_i}-W_i{}^{\mathrm{T}}\nabla\sigma_i\varepsilon_{T_i})\frac{\bar{B}_{i+N}^{\mathrm{T}}}{m_{s_{i+N}}}\tilde{W}_i - W_i{}^{\mathrm{T}}\nabla\sigma_i\times \\
& [\psi_i{}^{\star}\varsigma_ic_i(-c_iR_{ii}{}^{-1}(\psi_i{}^{\star}\varsigma_i)^{\mathrm{T}}\nabla\sigma_i{}^{\mathrm{T}}W_i - c_iR_{ii}{}^{-1}(\varepsilon_{g_i})^{\mathrm{T}}\times \\
& \nabla\sigma_i{}^{\mathrm{T}}W_i(c_iT_{ii}{}^{-1}(\beta_i{}^{\star}\vartheta_i)^{\mathrm{T}}\nabla\sigma_i{}^{\mathrm{T}}W_i + c_iT_{ii}{}^{-1}(\varepsilon_{k_i})^{\mathrm{T}}\times \\
& \nabla\sigma_i{}^{\mathrm{T}}W_i - c_iT_{ii}{}^{-1}((\tilde{\beta}_i+\beta_i{}^{\star})\vartheta_i)^{\mathrm{T}}\nabla\sigma_i{}^{\mathrm{T}}W_i \\
& +c_iT_{ii}{}^{-1}((\tilde{\beta}_i+\beta_i{}^{\star})\vartheta_i)^{\mathrm{T}}\nabla\sigma_i{}^{\mathrm{T}}\tilde{W}_{i+2N})\nabla\sigma_i{}^{\mathrm{T}}W_i \\
& -c_iT_{ii}{}^{-1}((\tilde{\beta}_i+\beta_i{}^{\star})\vartheta_i)^{\mathrm{T}}\nabla\sigma_i{}^{\mathrm{T}}W_i \; + c_iT_{ii}{}^{-1}\times \\
& ((\tilde{\beta}_i+\beta_i{}^{\star})\vartheta_i)^{\mathrm{T}}\nabla\sigma_i{}^{\mathrm{T}}\tilde{W}_{i+2N}) + \frac{1}{\gamma^2}\sum_{j\in N_i}e_{ij}\times
\end{aligned}$$

$$(\beta_j{}^{\star}\vartheta_j+\varepsilon_{k_j})(c_jT_{jj}{}^{-1}(\beta_j{}^{\star}\vartheta_j)^{\mathrm{T}}\nabla\sigma_j{}^{\mathrm{T}}W_j$$
$$+c_jT_{jj}{}^{-1}(\varepsilon_{k_j})^{\mathrm{T}}\nabla\sigma_j{}^{\mathrm{T}}W_j$$
$$-c_jT_{jj}{}^{-1}((\tilde{\beta}_j+\beta_j{}^{\star})\vartheta_j)^{\mathrm{T}}\nabla\sigma_j{}^{\mathrm{T}}W_j$$
$$+c_jT_{jj}{}^{-1}((\tilde{\beta}_j+\beta_j{}^{\star})\vartheta_j)^{\mathrm{T}}\nabla\sigma_j{}^{\mathrm{T}}\tilde{W}_{j+2N})]\frac{\bar{B}^{\mathrm{T}}_{i+N}}{m_{s_{i+N}}}\tilde{W}_i,$$

$$r_i=1-\frac{a_{i_2}}{2\varepsilon m_{s_{i+N}}},c_i=(d_i+e_{i0}),c_j=(d_j+e_{j0}),$$
$$\left\|W_i{}^{\mathrm{T}}\nabla\sigma_i(\sum_{j\in N_i}e_{ij}(\tilde{\theta}_i\xi_i-\tilde{\theta}_j\xi_j)+e_{i0}(\tilde{\theta}_i\xi_i-\tilde{\theta}_0\xi_0))\right\|\leqslant$$
$$a_{i_1}+a_{i_2}\|\delta_i\|,k_{T_i}=k_{\hat{D}_i}+\sum_{j\in N_i}k_{\hat{E}_j}+k_{\hat{H}_i}+\sum_{j\in N_i}k_{\hat{G}_j},$$
$$\left\|\frac{1}{2}W_i{}^{\mathrm{T}}D_iW_i\right\|\leqslant k_{D_i},\left\|\frac{1}{2}W_i{}^{\mathrm{T}}\hat{D}_iW_i\right\|\leqslant k_{\hat{D}_i},$$
$$\left\|\frac{1}{2}W_j{}^{\mathrm{T}}E_jW_j\right\|\leqslant k_{E_j},\left\|\frac{1}{2}W_j{}^{\mathrm{T}}\hat{E}_jW_j\right\|\leqslant k_{\hat{E}_j},$$
$$\left\|\frac{1}{2}W_i{}^{\mathrm{T}}H_iW_i\right\|\leqslant k_{H_i},\left\|\frac{1}{2}W_i{}^{\mathrm{T}}\hat{H}_iW_i\right\|\leqslant k_{\hat{H}_i},$$
$$\left\|\frac{1}{2}W_j{}^{\mathrm{T}}G_jW_j\right\|\leqslant k_{G_j},\left\|\frac{1}{2}W_j{}^{\mathrm{T}}\hat{G}_jW_j\right\|\leqslant k_{\hat{G}_j},$$
$$\left\|\hat{D}_i\right\|\leqslant k_{N_i},\left\|\hat{E}_j\right\|\leqslant k_{O_j},\left\|\hat{H}_i\right\|\leqslant k_{P_i},\left\|\hat{G}_j\right\|\leqslant k_{Q_j}.$$

Note that $k_{D_i},k_{\hat{D}_i},k_{E_j},k_{\hat{E}_j},k_{H_i},k_{\hat{H}_i},k_{G_j},k_{\hat{G}_j},k_{N_i},k_{O_j}$, k_{P_i} and k_{Q_j} are constant scalars. Using (39),(41),(16) and (34), (38) becomes

$$\dot{L}(t)\leqslant\sum_{i=1}^{N}\Big\{\dot{L}\_i+\dot{\bar{L}}_{V_i}+\tilde{W}^{\mathrm{T}}_{i+N}\hat{D}_i\tilde{W}_{i+N}\frac{\bar{B}^{\mathrm{T}}_{i+N}}{2m_{s_{i+N}}}W_i$$
$$-\tilde{W}^{\mathrm{T}}_{i+N}\hat{D}_iW_i\frac{\bar{B}^{\mathrm{T}}_{i+N}}{2m_{s_{i+N}}}W_i+\tilde{W}^{\mathrm{T}}_{i+N}\hat{D}_iW_i\times$$
$$\frac{\bar{B}^{\mathrm{T}}_{i+N}}{2m_{s_{i+N}}}\tilde{W}_i+\sum_{j\in N_i}\tilde{W}^{\mathrm{T}}_{j+N}\hat{E}_j\tilde{W}_{j+N}\frac{\bar{B}^{\mathrm{T}}_{i+N}}{2m_{s_{i+N}}}W_i$$
$$-\sum_{j\in N_i}\tilde{W}^{\mathrm{T}}_{j+N}\hat{E}_jW_j\frac{\bar{B}^{\mathrm{T}}_{i+N}}{2m_{s_{i+N}}}W_i+\sum_{j\in N_i}\tilde{W}^{\mathrm{T}}_{j+N}\hat{E}_j\times$$
$$W_j\frac{\bar{B}^{\mathrm{T}}_{i+N}}{2m_{s_{i+N}}}\tilde{W}_i-\frac{1}{\gamma^2}\tilde{W}^{\mathrm{T}}_{i+2N}\hat{H}_i\tilde{W}_{i+2N}\frac{\bar{B}^{\mathrm{T}}_{i+N}}{2m_{s_{i+N}}}W_i$$
$$+\frac{1}{\gamma^2}\tilde{W}^{\mathrm{T}}_{i+2N}\hat{H}_iW_i\frac{\bar{B}^{\mathrm{T}}_{i+N}}{2m_{s_{i+N}}}W_i-\frac{1}{\gamma^2}\tilde{W}^{\mathrm{T}}_{i+2N}\hat{H}_i\times$$
$$W_i\frac{\bar{B}^{\mathrm{T}}_{i+N}}{2m_{s_{i+N}}}\tilde{W}_i-\frac{1}{\gamma^2}\sum_{j\in N_i}\tilde{W}^{\mathrm{T}}_{j+2N}\hat{G}_j\tilde{W}_{j+2N}\frac{\bar{B}^{\mathrm{T}}_{i+N}}{2m_{s_{i+N}}}W_i$$
$$+\frac{1}{\gamma^2}\sum_{j\in N_i}\tilde{W}^{\mathrm{T}}_{j+2N}\hat{G}_jW_j\frac{\bar{B}^{\mathrm{T}}_{i+N}}{2m_{s_{i+N}}}W_i$$
$$-\frac{1}{\gamma^2}\sum_{j\in N_i}\tilde{W}^{\mathrm{T}}_{j+2N}\hat{G}_jW_j\frac{\bar{B}^{\mathrm{T}}_{i+N}}{2m_{s_{i+N}}}\tilde{W}_i\ \}$$
$$+\sum_{i=1}^{N}L_{W_{i+N}}+\sum_{i=1}^{N}L_{W_{i+2N}},\tag{42}$$

where,

$$\sum_{i=1}^{N}L_{W_{i+N}}=-\sum_{i=1}^{N}\tilde{W}^{\mathrm{T}}_{i+N}\alpha_{i+N}{}^{-1}\dot{\hat{W}}_{i+N}+\sum_{i=1}^{N}\tilde{W}^{\mathrm{T}}_{i+N}\hat{D}_i\times$$
$$\hat{W}_{i+N}\frac{\bar{B}^{\mathrm{T}}_{i+N}}{2m_{s_{i+N}}}\hat{W}_i+\underbrace{\sum_{i=1}^{N}\sum_{j\in i_N}\tilde{W}^{\mathrm{T}}_{j+N}\hat{E}_j\hat{W}_{j+N}\frac{\bar{B}^{\mathrm{T}}_{j+N}}{2m_{s_j}}\hat{W}_j}_{\sum_{i=1}^{N}\tilde{W}^{\mathrm{T}}_{i+N}\sum_{j\in i_{N_i}}\hat{E}_j\hat{W}_{i+N}\frac{\bar{B}^{\mathrm{T}}_{j+N}}{2m_{s_j}}\hat{W}_j}),$$

$$\sum_{i=1}^{N}L_{W_{i+2N}}=\frac{1}{\gamma^2}\sum_{i=1}^{N}\tilde{W}^{\mathrm{T}}_{i+2N}\alpha_{i+2N}{}^{-1}\dot{\hat{W}}_{i+2N}$$
$$-\frac{1}{\gamma^2}\sum_{i=1}^{N}\tilde{W}^{\mathrm{T}}_{i+2N}\hat{H}_i\hat{W}_{i+2N}\frac{\bar{B}^{\mathrm{T}}_{i+N}}{2m_{s_{i+N}}}\hat{W}_i$$
$$-\frac{1}{\gamma^2}\underbrace{\sum_{i=1}^{N}\sum_{j\in i_N}\tilde{W}^{\mathrm{T}}_{j+2N}\hat{G}_j\hat{W}_{j+2N}\frac{\bar{B}^{\mathrm{T}}_{j+N}}{2m_{s_j}}\hat{W}_j}_{\sum_{i=1}^{N}\tilde{W}^{\mathrm{T}}_{i+2N}\sum_{j\in i_{N_i}}\hat{G}_j\hat{W}_{i+2N}\frac{\bar{B}^{\mathrm{T}}_{j+N}}{2m_{s_j}}\hat{W}_j}).$$

Using $L_{W_{i+N}}$ and $L_{W_{i+2N}}$, then $\dot{\hat{W}}_{i+N}(t)$ and $\dot{\hat{W}}_{i+2N}(t)$ are obtained as defined in (35) and (36) respectively. Replacing (35) and (36), then $\dot{\hat{W}}_{i+N}(t)$ and $\dot{\hat{W}}_{i+2N}(t)$ are given,

$$\tilde{W}^{\mathrm{T}}_{i+N}S_iW_i-\tilde{W}^{\mathrm{T}}_{i+N}S_i\tilde{W}_{i+N}$$
$$-\tilde{W}^{\mathrm{T}}_{i+N}F_iW_i+\tilde{W}^{\mathrm{T}}_{i+N}F_i\tilde{W}_i\tag{43}$$

$$\tilde{W}^{\mathrm{T}}_{i+N}S_iW_i-\tilde{W}^{\mathrm{T}}_{i+N}S_i\tilde{W}_{i+N}$$
$$-\tilde{W}^{\mathrm{T}}_{i+N}F_iW_i+\tilde{W}^{\mathrm{T}}_{i+N}F_i\tilde{W}_i.\tag{44}$$

Since $Q_i(\delta_i)>0,\ i=1,\cdots,N$, there exists $q_i>0,\ \forall i$ such that $-\delta_i{}^{\mathrm{T}}q_i\delta_i>-\delta_i{}^{\mathrm{T}}Q_i\delta_i$. Now $\dot{L}$ can be reformed as

$$\dot{L}(t)=\sum_{i=1}^{N}\{C_{ii}-\tilde{Z}_i^{\mathrm{T}}M_i\tilde{Z}_i+\ D_{ii}\tilde{Z}_i\ \},$$
$$\tilde{Z}_i=\left[\delta_i,\tilde{W}_i,\tilde{W}_{i+N},\tilde{W}_{j+N},\tilde{W}_{i+2N},\tilde{W}_{j+2N}\right]^{\mathrm{T}},\tag{45}$$

where $C_{ii}\leqslant C_{i\,\max},\ \ D_{ii}\leqslant D_{i\,\max}$. Let the parameters $S_i,\bar{S}_i,F_i$ and $\bar{F}_i$ be chosen such that the squared matrix M_i is positive definite. Finally (42) becomes

$$\dot{L}<\sum_{i=1}^{N}\{-\left\|\tilde{Z}_i\right\|^2\sigma_{\min}(M_i)+D_{i\,\max}\left\|\tilde{Z}_i\right\|+C_{i\,\max}\}.\tag{46}$$

The Lyapunov derivative is negative as long as

$$\left\|\tilde{Z}_i\right\|>$$
$$\frac{D_{i\,\max}}{2\sigma_{\min}(M_i)}+\sqrt{\frac{D_{i\,\max}{}^2}{4\sigma_{\min}{}^2(M_i)}\left\|\tilde{Z}_i\right\|+\frac{C_{i\,\max}}{\sigma_{\min}(M_i)}}.\tag{47}$$

According to [31] we can show that if (47) exceeds a certain bound, then $\dot{L}$ is negative and the closed-loop signals are UUB. ■