Automatica 185 (2026) 112816

Contents lists available at ScienceDirect

Automatica

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

# Distributed Nash equilibrium computation in multi-group resource allocation games over digraphs[☆]

Jialing Zhou[a], Guanghui Wen[b,*], Yuezu Lv[a], Xinlei Yi[c], Tao Yang[d], Karl Henrik Johansson[e]

[a] State Key Laboratory of Environment Characteristics and Effects for Near-space, Beijing Institute of Technology, Beijing 100081, China
[b] School of Automation, Southeast University, Nanjing 210096, China
[c] Shanghai Institute of Intelligent Science and Technology, Tongji University, Shanghai 201210, China
[d] State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110004, China
[e] School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, SE-100 44, Stockholm, Sweden

ARTICLE INFO

*Article history:*
Received 17 October 2023
Received in revised form 21 July 2025
Accepted 2 December 2025
Available online 2 January 2026

*Keywords:*
Distributed resource allocation
Directed graph
Hard balance constraint
Nash equilibrium seeking
Resource allocation game

ABSTRACT

The existing distributed resource allocation (DRA) algorithms for multi-agent networks can rarely be implemented for multiple interacting groups of agents with conflicts of interest. The directed interaction, together with the hard balance constraint that follows from maintaining supply–demand balance during the execution process, make the DRA more challenging. To address this problem, the paper studies DRA over multiple interacting groups from a game-theoretic perspective, introducing the resource allocation game (RAG). A novel out-Laplacian matrix based methodology is developed for distributed Nash equilibrium (NE) computation. Following this methodology, distributed algorithms are designed using leader-follower tracking protocols to estimate partial derivatives of individual objective functions for the RAG. A reduced-order distributed algorithm is further developed for the RAG by integrating a gradient-tracking mechanism for estimating partial derivatives of group-level objective functions. It is shown that agent states converge to the NE of the games linearly while satisfying the balance constraint during the whole execution process under the proposed algorithms. The effectiveness of the proposed algorithms is illustrated through numerical examples.


## 1. Introduction

The current trend in resource allocation over multi-agent networks demands the design of decision-making using only local information. The problem is usually referred to as distributed resource allocation (DRA), with the standard form of optimizing a group-level objective under coupled equality constraints and partial information setting (Xiao & Boyd, 2006). Many distributed optimization algorithms, such as the initialization-free (Yi, Hong, & Liu, 2016), stochastic approximation (Yi, Lei, & Hong, 2018), Newton-like (Anderson & Martínez, 2021), linear-quadratic saddle-point (Simpson-Porco, Poolla, Monshizadeh, & Dörfler, 2020), and distributed multi-proximal primal–dual (Wei, Shang, Fang, Zeng, Dou et al., 2022) algorithms, have been developed for the standard DRA problem. For a recent overview, see Yang et al. (2019).

In some situations, such as economic dispatch in energy systems, it is preferable to meet *hard balance constraints*, to maintain supply–demand balances during the process of finding the optimal solution in an online implementation. Most existing works consider the balance demand only at the settlement of the DRA algorithm, i.e., soft balance constraints; DRA under hard balance constraints is investigated in Chen and Li (2018) and Zhou, Lv, Wen, and Wen (2022) for undirected and directed multi-agent networks, respectively.

The aforementioned literature primarily focuses on DRA within a single group of cooperative agents. However, in many practical scenarios, the benefit received by each group depends not only on its internal resource allocation but also on external market factors, such as the total supply of similar services offered by other groups. Although resources are allocated internally, the overall group-level performance is coupled through such shared environments, introducing conflicts of interest among groups and rendering single-group DRA approaches inadequate. Such scenarios can be captured by the multi-group resource allocation game (RAG) formulation, which reflects both intra-group cooperation and inter-group competition. Specifically, each group of agents

☆ This work was supported by the National Natural Science Foundation of China under Grant Nos. 62376029, 62325304, U24A20279, 62522307, 62273045, U2341213, 62088101, and 62503365. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Sergio Grammatico under the direction of Editor Florian Dorfler.
* Corresponding author.
*E-mail addresses:* jlzhou@bit.edu.cn (J. Zhou), ghwen@seu.edu.cn (G. Wen), yzlv@bit.edu.cn (Y. Lv), xinleiyi@tongji.edu.cn (X. Yi), yangtao@mail.neu.edu.cn (T. Yang), kallej@kth.se (K.H. Johansson).

https://doi.org/10.1016/j.automatica.2025.112816
0005-1098/

seeks to allocate a shared group-level resource to maximize its overall utility, while accounting for the fact that its benefit may be directly influenced by the resource allocation strategies of other groups. This coupling across groups transforms the DRA problem into a game-theoretic setting and requires new algorithmic tools. A natural solution concept in this context is the Nash equilibrium (NE). Recent progress has been reported on distributed NE computation in multi-group games (Lou, Hong, Xie, Shi, & Johansson, 2016; Meng & Li, 2023; Nian, Niu, & Yang, 2022; Pang & Hu, 2022; Zeng, Chen, Liang, & Hong, 2019; Zhou, Lv, Wen, Lü, & Zheng, 2023), where the agents within each group are either not subject to any coupled constraints, or subject to a special kind of coupled constraint that requires the individual agent states to be equal. Few attempts have been made to address multi-group games subject to resource allocation constraints, as noted in Deng and Liu (2023) and Zhou, Wen, Lv, Yang, and Chen (2024). The algorithms proposed in these works require that the Laplacian matrices of the group-level subgraphs have a left eigenvector of **1** associated with zero eigenvalue, making them inapplicable to general directed graphs. Designing distributed NE computation algorithms for RAG with unbalanced directed topologies presents significant challenges. The hard balance constraint further complicates this task.

Motivated by the observations above, this paper studies the RAG subject to the hard balance constraint over directed network topologies. To meet the hard balance constraint, a novel out-Laplacian matrix based strategy for distributed NE computation is proposed, where the out-Laplacian matrix acts as a singular linear transformation to build the relation between the agent state and a virtual variable, making it possible to design the virtual variable iterations to achieve successful NE computation also under general directed network topologies. Based on this strategy, two DRA algorithms are designed with different mechanisms of estimating the required partial derivatives of objective functions. First, a DRA algorithm is proposed by employing the leader-follower tracking method for the estimation of individual agents states as well as partial derivatives of individual objective functions. To reduce the computational complexity, a new kind of reduced-order DRA algorithm is then designed, where a gradient-tracking law is integrated in the estimation of the partial derivatives of the group-level objective functions rather than the individual objective functions.

The main contributions of the paper are summarized as follows. (i) A new out-Laplacian matrix based distributed NE computation design framework is developed for the RAG, which successfully overcomes the difficulties caused by the hard balance constraint and directed network topologies. One distinct feature of the proposed approach as compared to primal–dual approaches is that supply–demand balance can be maintained at every iteration, so as to facilitate feasible online implementation. (ii) Two novel types of DRA algorithms are developed using the out-Laplacian framework. One is designed by integrating a leader-follower consensus tracking mechanism, which is convenient for convergence analysis ; while the other is developed by integrating a gradient-tracking mechanism, which reduces the computational complexity. (iii) Linear convergence of the collective state to the exact NE of the RAG is established under the proposed DRA algorithms, requiring the development of novel analysis techniques.

The remainder of the paper is organized as follows. Section 2 provides the problem statement. In Section 3, a novel DRA algorithm is developed and the convergence analysis is presented. A reduced-order DRA algorithm is further designed and analyzed in Section 4. Section 5 shows numerical simulation results, and Section 6 concludes the paper.

**Notations and Definitions**. Symbols $\mathbb{R}$, $\mathbb{N}$, and $\mathbb{N}^+$, respectively, denote the sets of real numbers, natural numbers, and positive integers. The set of $n$-dimensional real column vectors is denoted by $\mathbb{R}^n$. $I_n$ is the $n$-dimensional identity matrix. $\mathbf{1}_n(\mathbf{0}_n)$ denotes the $n$-dimensional column vector with all the entries being 1(0). Symbols $\otimes$ and $\|\cdot\|$ denote the Kronecker product and the Euclidean norm, respectively. $(\cdot)_i$ denotes the $i$th element of a vector or the $i$th row of a matrix. The expression $\mathrm{col}(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$ denotes a column vector that is formed by vertically stacking the given column vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ together. For matrices $A_1, A_2, \ldots, A_n$ of arbitrary dimensions, $\mathrm{diag}\{A_1, A_2, \ldots, A_n\}$ denotes the diagonal block matrix with $A_1, A_2, \ldots, A_n$ being the diagonal blocks. For a symmetric matrix $A$, $\lambda_2(A)$ denotes the second smallest eigenvalue.

## 2. Problem statement

A RAG among $N \in \mathbb{N}^+$ groups of agents under general directed network topologies is considered. Denote the set of group indices by $\mathcal{I} = \{1, \ldots, N\}$. Denote the set of agents in group $i \in \mathcal{I}$ by $\mathcal{V}_i = \{i1, i2 \ldots, in_i\}$, where $il$ represents the agent $l$ in group $i$ and $n_i = |\mathcal{V}_i|$. Define the overall set of the game participants $\mathcal{V} = \bigcup_{i\in\mathcal{I}} \mathcal{V}_i$, and their number $n_{\mathrm{sum}} = |\mathcal{V}|$.

### 2.1. Communication topology and associated matrices

The communication topology among the $n_{\mathrm{sum}}$ agents is described by a directed graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ with the node (agent) set $\mathcal{V}$ and the edge (communication link) set $\mathcal{E} \subseteq \mathcal{V}\times\mathcal{V}$. A pair $(ij, lm) \in \mathcal{E}$ is an edge of $\mathcal{G}$ if agent $lm$ can receive information from agent $ij$. If $(ij, lm) \in \mathcal{E}$, then agent $ij$ is called an *in-neighbor* of agent $lm$, and agent $lm$ is called an *out-neighbor* of agent $ij$. For each agent $ij \in \mathcal{V}$, define the *in-neighbor set* $\mathcal{N}_{ij} = \{pq|(pq, ij) \in \mathcal{E}\}$, the *in-degree* $d_{ij} = |\mathcal{N}_{ij}|$, the *out-neighbor set* $\mathcal{N}^{\mathrm{o}}_{ij} = \{pq|(ij, pq) \in \mathcal{E}\}$, and the *out-degree* $d^{\mathrm{o}}_{ij} = |\mathcal{N}^{\mathrm{o}}_{ij}|$. A directed path from agent $i_1j_1$ to agent $i_lj_l$ is a sequence of edges $(i_mj_m, i_{m+1}j_{m+1}) \in \mathcal{E}, m = 1, \ldots, l-1$. A graph is called strongly connected if there exist directed paths from every node to every other node.

The communication graph among the agents inside group $i\in\mathcal{I}$, denoted $\mathcal{G}_i(\mathcal{V}_i, \mathcal{E}_i)$, can be derived from graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$. Obviously, $\mathcal{E}_i = \{(ij, il)|(ij, il)\in\mathcal{E}\}$. For each agent $ij\in\mathcal{V}$, define the *intra-group in-neighbor set* $\mathcal{N}^{i}_{ij} = \{im|(im, ij) \in \mathcal{E}_i\}$, the *intra-group in-degree* $d^{i}_{ij} = |\mathcal{N}^{i}_{ij}|$, the *intra-group out-neighbor set* $\mathcal{N}^{i\mathrm{o}}_{ij} = \{im|(ij, im) \in \mathcal{E}_i\}$ and the *intra-group out-degree* $d^{i\mathrm{o}}_{ij} = |\mathcal{N}^{i\mathrm{o}}_{ij}|$.

**Assumption 1.** The graph $\mathcal{G}$ and all the induced subgraphs $\mathcal{G}_1, \ldots, \mathcal{G}_N$ are strongly connected.

In the following, several $n_{\mathrm{sum}}$-dimensional square matrices associated with graph $\mathcal{G}$ are introduced. The superscript and subscript in the entries of these matrices $(\cdot)^{pq}_{ij}$, denote that the element is on the $(\sum_{k=1}^{i-1} n_k+j)$th row and the $(\sum_{k=1}^{p-1} n_k+q)$th column.

- $\mathcal{A} = [a^{pq}_{ij}]_{n_{\mathrm{sum}}\times n_{\mathrm{sum}}}$ is the adjacency matrix of $\mathcal{G}$ with $a^{ij}_{ij} = 0$, and $a^{pq}_{ij} = 1$ if $(pq, ij) \in \mathcal{E}$ and 0 otherwise.
- $\mathcal{W} = [w^{pq}_{ij}]_{n_{\mathrm{sum}}\times n_{\mathrm{sum}}}$ is the weighted adjacency matrix of $\mathcal{G}$ with $w^{ij}_{ij} = 0$, $w^{pq}_{ij} > 0$ if $(pq, ij) \in \mathcal{E}$ and 0 otherwise. Moreover, the entries satisfy $\sum_{pq\in\mathcal{V}} w^{pq}_{ij} + \max_{pq\in\mathcal{V}}\{w^{pq}_{ij}\} < 1$, $\forall ij \in \mathcal{V}$. For example, the entries can be set as $w^{pq}_{ij} = a^{pq}_{ij}/(d_{ij} + h_{ij})$, where $h_{ij} > \max_{pq\in\mathcal{V}}\{a^{pq}_{ij}\}$ is a constant.

Similarly, several $n_i$-dimensional square matrices associated with graph $\mathcal{G}_i$ $\forall i \in \mathcal{I}$ are defined. The superscript and subscript in the entries of these matrices $(\cdot)^{il}_{ij}$, denote that the element is on the $j$th row and the $l$th column.

- $\mathcal{A}_i = [a_{ij}^{il}]_{n_i \times n_i}$ is the adjacency matrix of $\mathcal{G}_i$ with the entries defined consistent with those in $\mathcal{A}$. Obviously, $\mathcal{A}_1, \ldots, \mathcal{A}_N$ are the diagonal blocks of $\mathcal{A}$.
- $\mathcal{L}_i^{\mathrm{o}} = [l_{ij}^{il}]_{n_i \times n_i}$ is the out-Laplacian matrix of $\mathcal{G}_i$, defined as $l_{ij}^{ij} = \sum_{m=1}^{n_i} a_{im}^{ij}$ and $l_{ij}^{il} = -a_{ij}^{il}$, $j \neq l$. Obviously, $\mathcal{L}_i^{\mathrm{o}} = \mathcal{D}_i^{\mathrm{o}} - \mathcal{A}_i$, where $\mathcal{D}_i^{\mathrm{o}} = \operatorname{diag}\{d_{i1}^{\mathrm{io}}, \ldots, d_{in_i}^{\mathrm{io}}\}$.
- $\mathcal{R}_i = [r_{ij}^{il}]_{n_i \times n_i}$ is the weighted adjacency matrix of $\mathcal{G}_i$ with $r_{ij}^{ij} = 0$, $r_{ij}^{il} > 0$ if $(il, ij) \in \mathcal{E}_i$ and 0 otherwise. Moreover, the entries satisfy $\sum_{il \in \mathcal{V}_i} r_{ij}^{il} + \max_{il \in \mathcal{V}_i}\{r_{ij}^{il}\} < 1$, $\forall ij \in \mathcal{V}_i$. For example, the entries can be set as $r_{ij}^{il} = a_{ij}^{il}/(d_{ij}^{i} + h_{ij}^{i})$, where $h_{ij}^{i} > \max_{il \in \mathcal{V}_i}\{a_{ij}^{il}\}$ is a constant.
- $\mathcal{C}_i = [c_{ij}^{il}]_{n_i \times n_i}$ is a column-stochastic adjacency matrix associated with $\mathcal{G}_i$, with $c_{ij}^{il} > 0$ if $ij \in \mathcal{N}_{il}^{\mathrm{io}} \cup \{il\}$ and 0 otherwise. For example, the entries can be set as $c_{ij}^{il} = a_{ij}^{il}/(1 + d_{il}^{\mathrm{o}})$, $\forall j \neq l$, and $c_{il}^{il} = 1/(1 + d_{il}^{\mathrm{o}})$.

### 2.2. RAG

Each agent $ij$ has a certain amount of local resources, denoted $R_{ij} \in \mathbb{R}$. In each group $i$, the members collaborate to make the best decision on the re-allocation of the group resources, represented by $R_i = \sum_{ij \in \mathcal{V}_i} R_{ij}$, for the maximization of group-level benefit, i.e., the sum of individual benefits of the group members.

Define $x_{ij} \in \mathbb{R}$ the state of agent $ij$, representing the decision on the quantity of resources re-allocated to agent $ij$. Denote the state of group $i$ by $\boldsymbol{x}_i = \operatorname{col}(x_{i1}, x_{i2}, \ldots, x_{in_i}) \in \mathbb{R}^{n_i}$, and define $\boldsymbol{x} = \operatorname{col}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N) \in \mathbb{R}^{n_{\mathrm{sum}}}$ and $\boldsymbol{x}_{-i} = \operatorname{col}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{i-1}, \boldsymbol{x}_{i+1}, \ldots, \boldsymbol{x}_N)$. Denote the objective function of group $i$ by $f_i$, and the private individual objective function of agent $ij$ by $f_{ij}$.

In a complex competitive environment, the benefit of each group may be influenced by the states of agents not only inside but also outside the group. The conflicts of interest among the groups make the resource allocation problem a multi-group game. In this paper, the following form of RAG is studied.

$$
\begin{aligned}
&\min_{\boldsymbol{x}_i} f_i(\boldsymbol{x}) = \min_{\boldsymbol{x}_i} \sum_{il \in \mathcal{V}_i} f_{il}(\boldsymbol{x}), \ \forall i \in \mathcal{I}, \\
&\text{s.t.} \quad \sum_{ij \in \mathcal{V}_i} x_{ij} = R_i.
\end{aligned} \tag{1}
$$

**Example 1.** Consider an internet-based company operating multiple business lines (e.g., cloud storage, video streaming, search services), each of which requires computing power to deliver its services. The company aims to allocate its limited computing resources among these business lines to maximize overall revenue. The service capacity of each business line depends on the amount of computing power it receives. Meanwhile, the market price or revenue per unit of service is influenced by the total supply of similar services across all companies. As a result, the revenue of each business line is affected not only by the internal resource allocation within its own company, but also by the resource decisions made by competing business lines from other companies. This scenario corresponds to the RAG model in (1), where each company is modeled as a group, its business lines as agents, and the group-level objective is coupled with those of other groups through the shared market environment. For a detailed numerical illustration of this example, please refer to the simulation case in Section 5.

Define $\Omega_i = \{\boldsymbol{x}_i \in \mathbb{R}^{n_i} | \mathbf{1}_{n_i}^{\mathrm{T}} \boldsymbol{x}_i = R_i\}$ and $\Omega = \prod_{i=1}^{N} \Omega_i$. Then, the definition of the NE of RAG can be given as follows.

**Definition 1.** A vector $\boldsymbol{x}^* = \operatorname{col}(\boldsymbol{x}_1^*, \ldots, \boldsymbol{x}_N^*) \in \Omega$ is an NE of RAG (1) if $\forall \boldsymbol{x}_i \in \Omega_i$, $\forall i \in \mathcal{I}$:

$$
\begin{aligned}
&f_i\left(\operatorname{col}(\boldsymbol{x}_1^*, \ldots, \boldsymbol{x}_{i-1}^*, \boldsymbol{x}_i^*, \boldsymbol{x}_{i+1}^*, \ldots, \boldsymbol{x}_N^*)\right) \\
\leq &f_i\left(\operatorname{col}(\boldsymbol{x}_1^*, \ldots, \boldsymbol{x}_{i-1}^*, \boldsymbol{x}_i, \boldsymbol{x}_{i+1}^*, \ldots, \boldsymbol{x}_N^*)\right).
\end{aligned}
$$

Define the pseudo gradient $\mathbf{P} : \mathbb{R}^{n_{\mathrm{sum}}} \to \mathbb{R}^{n_{\mathrm{sum}}}$ as

$$
\mathbf{P}(\cdot) = \operatorname{col}\left(\frac{\partial f_1}{\partial \boldsymbol{x}_1}(\cdot), \frac{\partial f_2}{\partial \boldsymbol{x}_2}(\cdot), \ldots, \frac{\partial f_N}{\partial \boldsymbol{x}_N}(\cdot)\right).
$$

**Assumption 2.** For each agent $ij \in \mathcal{V}$ in RAG (1), the objective function $f_{ij}$ is convex and continuously differentiable in $\boldsymbol{x}_i$ for each $\boldsymbol{x}_{-i} \in \mathbb{R}^{n_{\mathrm{sum}} - n_i}$. Moreover, $\frac{\partial f_{ij}}{\partial \boldsymbol{x}_i}(\cdot)$ is Lipschitz with the constant $l_{ij}$.

**Assumption 3.** There exists a positive constant $\mu$ such that $(\boldsymbol{a}_1 - \boldsymbol{a}_2)^{\mathrm{T}}(\mathbf{P}(\boldsymbol{a}_1) - \mathbf{P}(\boldsymbol{a}_2)) \geq \mu \|\boldsymbol{a}_1 - \boldsymbol{a}_2\|^2$, $\forall \boldsymbol{a}_1, \boldsymbol{a}_2 \in \Omega$.

The above two assumptions ensure the existence and uniqueness of the Nash equilibrium (NE) of the considered RAG problem. They are aligned with the assumptions used in studies of multi-group networked games (Meng & Li, 2023; Pang & Hu, 2022), and are also consistent with the assumptions in other classes of networked games (Huang, Lei, & Hong, 2023). It is worth noting that Assumption 3 in this work slightly relaxes the standard condition by requiring the pseudo-gradient mapping $\mathbf{P}$ to be strongly monotone only over the feasible set $\Omega$ (defined by resource constraints), rather than over the full decision space as is commonly assumed in the existing literature. In addition, Assumption 2 directly implies that each group-level gradient $\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\cdot)$ is $l_i$-Lipschitz continuous, where $l_i = \sum_{j=1}^{n_i} l_{ij}$.

**Lemma 1.** *Suppose Assumptions 1 and 2 hold. A vector $\boldsymbol{x}^* \in \Omega$ is the NE of the RAG if and only if $\boldsymbol{x}^* \in \Omega$ satisfies $\mathcal{L}_i^{\mathrm{o}\mathrm{T}} \frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}^*) = \mathbf{0}_{n_i}$, $\forall i \in \mathcal{I}$.*

The proof of this lemma is given in Appendix A.

### 2.3. Design objective and constraint

We first introduce the following definition of *hard balance constraint.*

**Definition 2.** Hard balance constraint refers to the requirement of maintaining balance for all iterations $k$ in the computation process, i.e., $\sum_{ij \in \mathcal{V}_i} x_{ij}(k) = R_i$, $\forall k \in \mathbb{N}$.

The model with the equality constraint is of practical significance, particularly for scenarios requiring complete resource utilization or precise supply–demand balance in distributed economic dispatch problems. For online implementation in distributed economic dispatch problems, it is often desired to meet the supply–demand balance at every iteration, thus necessitating the hard balance constraint.

In this paper, we investigate the RAG problem within a *partial information setting.* Here, each agent $ij$ has access to only two types of information: its own private information — including its individual objective function $f_{ij}$ and state $x_{ij}$ — and the information received from its direct neighbors via communication links. Note that each agent $ij$ does not have full knowledge of the group-level objective $f_i$ or the group state $\boldsymbol{x}_i$. This setting reflects a distributed decision-making framework, where each agent determines its own resource usage based solely on local information and neighbor interactions, without a global view.

The objective of this paper is to develop DRA algorithms to drive the agent states to converge to the NE of the RAG under the *partial information setting*, while meeting the hard balance constraint.

## 3. DRA algorithm for RAG

In this section, a novel distributed algorithm is designed for NE computation in the RAG. Linear convergence is established.

### 3.1. Algorithm design

To compute the NE of the RAG under the partial information setting and the hard balance constraint, a new out-Laplacian matrix based DRA algorithm is developed. In Algorithm 1, each agent $ij$ updates state $x_{ij}$, auxiliary scalar variables $z_{ij}, s_{ij}^{11}, \dots, s_{ij}^{Nn_N}$ and auxiliary $n_i$-dimensional column vector variables $\boldsymbol{\eta}_{ij}^{i1}, \dots, \boldsymbol{\eta}_{ij}^{in_i}$. Define vectors $\boldsymbol{s}_{ij}^{l} = \mathrm{col}(s_{ij}^{l1}, \dots, s_{ij}^{ln_l}) \in \mathbb{R}^{n_l}$ and $\boldsymbol{s}_{ij} = \mathrm{col}(s_{ij}^{11}, \dots, s_{ij}^{Nn_N}) \in \mathbb{R}^{n_{\mathrm{sum}}}$ for notational brevity. The iteration of the local variables of agent $ij$ is based on one-hop information from its neighbors and the local information:

$$x_{ij}(k+1) = x_{ij}(0) - d_{ij}^{\mathrm{io}} z_{ij}(k+1) + \sum_{im\in\mathcal{N}_{ij}^{i}} z_{im}(k+1), \tag{2a}$$

$$z_{ij}(k+1) = z_{ij}(k) + \alpha \sum_{im\in\mathcal{N}_{ij}^{\mathrm{io}}} \sum_{il\in\mathcal{V}_i} \Big( (\boldsymbol{\eta}_{ij}^{il})_j(k) - (\boldsymbol{\eta}_{ij}^{il})_m(k) \Big), \tag{2b}$$

$$\boldsymbol{\eta}_{ij}^{il}(k+1) = \bar{r}_{ij}^{il} \boldsymbol{\eta}_{ij}^{il}(k) + \sum_{im\in\mathcal{N}_{ij}^{i}} r_{ij}^{im} \boldsymbol{\eta}_{im}^{il}(k) + r_{ij}^{il} \frac{\partial f_{il}}{\partial \boldsymbol{x}_i}(\boldsymbol{s}_{il}(k)), \qquad \forall il \in \mathcal{V}_i, \tag{2c}$$

$$s_{ij}^{pq}(k+1) = \bar{w}_{ij}^{pq} s_{ij}^{pq}(k) + \sum_{lm\in\mathcal{N}_{ij}} w_{ij}^{lm} s_{lm}^{pq}(k) + w_{ij}^{pq} x_{pq}(k), \qquad \forall pq \in \mathcal{V}, \tag{2d}$$

where $(\boldsymbol{\eta}_{ij}^{il})_j$ denotes the $j$th entry of the column vector $\boldsymbol{\eta}_{ij}^{il}$, $\alpha$ is a small constant step size to be determined later, $d_{ij}^{\mathrm{io}}$, $r_{ij}^{il}$ $(\forall il \in \mathcal{V}_i)$, $w_{ij}^{pq}$ $(\forall pq \in \mathcal{V})$ are constant parameters associated with graphs as defined in Section 2.1, and

$$\begin{aligned} \bar{r}_{ij}^{il} &= 1 - \sum_{im\in\mathcal{N}_{ij}^{i}} r_{ij}^{im} - r_{ij}^{il}, \\ \bar{w}_{ij}^{pq} &= 1 - \sum_{lm\in\mathcal{N}_{ij}} w_{ij}^{lm} - w_{ij}^{pq}. \end{aligned} \tag{3}$$

Define column vectors $\boldsymbol{z}_i = \mathrm{col}(z_{i1}, \dots, z_{in_i}) \in \mathbb{R}^{n_i}$, $\boldsymbol{\eta}_{ij} = \mathrm{col}\left(\boldsymbol{\eta}_{ij}^{i1}, \dots, \boldsymbol{\eta}_{ij}^{in_i}\right) \in \mathbb{R}^{n_i^2}$, $\boldsymbol{\eta}_i = \mathrm{col}\left(\boldsymbol{\eta}_{i1}, \dots, \boldsymbol{\eta}_{in_i}\right) \in \mathbb{R}^{n_i^3}$, $\boldsymbol{s}_i = \mathrm{col}(\boldsymbol{s}_{i1}, \dots, \boldsymbol{s}_{in_i}) \in \mathbb{R}^{n_i n_{\mathrm{sum}}}$, $\boldsymbol{s} = \mathrm{col}(\boldsymbol{s}_1, \dots, \boldsymbol{s}_N) \in \mathbb{R}^{n_{\mathrm{sum}}^2}$, and the function $\mathbf{Q}_i : \mathbb{R}^{n_i n_{\mathrm{sum}}} \to \mathbb{R}^{n_i^2}$:

$$\begin{aligned} \mathbf{Q}_i(\boldsymbol{s}_i) = \mathrm{col}\bigg( & \frac{\partial f_{i1}}{\partial x_{i1}}(\boldsymbol{s}_{i1}), \frac{\partial f_{i1}}{\partial x_{i2}}(\boldsymbol{s}_{i1}), \dots, \frac{\partial f_{i1}}{\partial x_{in_i}}(\boldsymbol{s}_{i1}), \\ & \frac{\partial f_{i2}}{\partial x_{i1}}(\boldsymbol{s}_{i2}), \dots, \frac{\partial f_{i2}}{\partial x_{in_i}}(\boldsymbol{s}_{i2}), \dots, \frac{\partial f_{in_i}}{\partial x_{in_i}}(\boldsymbol{s}_{in_i}) \bigg). \end{aligned}$$

Then, the updating formulas in Algorithm 1 can be rewritten collectively for each group $i \in \mathcal{I}$ as follows:

$$\boldsymbol{x}_i(k+1) = \boldsymbol{x}_i(0) - \mathcal{L}_i^{\mathrm{o}} \boldsymbol{z}_i(k+1), \tag{4a}$$

$$\boldsymbol{z}_i(k+1) = \boldsymbol{z}_i(k) + \alpha \check{\mathcal{L}}_i^{\mathrm{o}} (I_{n_i} \otimes (\mathbf{1}_{n_i}^{\mathrm{T}} \otimes I_{n_i})) \boldsymbol{\eta}_i(k), \tag{4b}$$

$$\boldsymbol{\eta}_i(k+1) = ((\mathcal{R}_i \otimes I_{n_i} + \bar{\mathcal{R}}_i) \otimes I_{n_i}) \boldsymbol{\eta}_i(k) + (\hat{\mathcal{R}}_i \otimes I_{n_i})(\mathbf{1}_{n_i} \otimes \mathbf{Q}_i(\boldsymbol{s}_i(k))), \tag{4c}$$

$$\boldsymbol{s}(k+1) = (\mathcal{W} \otimes I_{n_{\mathrm{sum}}} + \bar{\mathcal{W}}) \boldsymbol{s}(k) + \hat{\mathcal{W}} \big(\mathbf{1}_{n_{\mathrm{sum}}} \otimes \boldsymbol{x}(k)\big), \tag{4d}$$

**Algorithm 1** DRA algorithm for the RAG

For agent $ij \in \mathcal{V}$

**Input:** intra-group out-degree $d_{ij}^{\mathrm{io}}$, weights $r_{ij}^{il}, \forall il \in \mathcal{V}_i$ and $w_{ij}^{pq}, \forall pq \in \mathcal{V}$ defined in Section 2.1, weights $\bar{r}_{ij}^{il}, \forall il \in \mathcal{V}_i$ and $\bar{w}_{ij}^{pq}, \forall pq \in \mathcal{V}$ defined in (3), and parameter $\alpha$.

**Initialize:** $x_{ij}(0) = R_{ij}$, $z_{ij}(0) = 0$, arbitrary $\boldsymbol{\eta}_{ij}^{il}(0) \in \mathbb{R}^{n_i}, \forall il \in \mathcal{V}_i$, and arbitrary $s_{ij}^{pq}(0) \in \mathbb{R}, \forall pq \in \mathcal{V}$.

**for** $k = 0, 1, 2, \dots$ **do**

  Compute $\frac{\partial f_{ij}}{\partial \boldsymbol{x}_i}(\boldsymbol{s}_{ij}(k))$.

  Send $x_{ij}(k)$, $\boldsymbol{s}_{ij}(k)$ to its out-neighbor $uv$ for all $uv \in \mathcal{N}_{ij}^{\mathrm{o}}$, and $z_{ij}(k)$, $\boldsymbol{\eta}_{ij}(k)$, $\frac{\partial f_{ij}}{\partial \boldsymbol{x}_i}(\boldsymbol{s}_{ij}(k))$ to its intra-group out-neighbor $ih$ for all $ih \in \mathcal{N}_{ij}^{\mathrm{io}}$. Receive $x_{pq}(k)$, $\boldsymbol{s}_{pq}(k)$ from its in-neighbor $pq$ for all $pq \in \mathcal{N}_{ij}$, and $z_{im}(k)$, $\boldsymbol{\eta}_{im}(k)$, $\frac{\partial f_{im}}{\partial \boldsymbol{x}_i}(\boldsymbol{s}_{im}(k))$ from its intra-group in-neighbor $im$ for all $im \in \mathcal{N}_{ij}^{i}$.

  Compute $z_{ij}(k+1)$ according to (2b).

  **for** $il \in \mathcal{V}_i$ **do**

    Compute $\boldsymbol{\eta}_{ij}^{il}(k+1)$ according to (2c).

  **end for**

  **for** $pq \in \mathcal{V}$ **do**

    Compute $s_{ij}^{pq}(k+1)$ according to (2d).

  **end for**

  Compute $x_{ij}(k+1)$ according to (2a).

**end for**

where the matrices $\mathcal{L}_i^{\mathrm{o}}$, $\mathcal{R}_i$ and $\mathcal{W}$ are defined in Section 2.1, $\check{\mathcal{L}}_i^{\mathrm{o}} = \mathrm{diag}\{(\mathcal{L}_i^{\mathrm{oT}})_1, \dots, (\mathcal{L}_i^{\mathrm{oT}})_{n_i}\} \in \mathbb{R}^{n_i \times n_i^2}$ with $(\mathcal{L}_i^{\mathrm{oT}})_j$ denoting the $j$th row of the matrix $\mathcal{L}_i^{\mathrm{oT}}$, $\bar{\mathcal{R}}_i = \mathrm{diag}\{\bar{r}_{i1}^{i1}, \dots, \bar{r}_{i1}^{in_i}, \bar{r}_{i2}^{i1}, \dots, \bar{r}_{in_i}^{in_i}\} \in \mathbb{R}^{n_i^2 \times n_i^2}$, $\hat{\mathcal{R}}_i = \mathrm{diag}\{r_{i1}^{i1}, \dots, r_{i1}^{in_i}, r_{i2}^{i1}, \dots, r_{in_i}^{in_i}\} \in \mathbb{R}^{n_i^2 \times n_i^2}$, $\bar{\mathcal{W}} = \mathrm{diag}\{\bar{w}_{11}^{11}, \dots, \bar{w}_{11}^{Nn_N}, \bar{w}_{12}^{11}, \dots, \bar{w}_{12}^{Nn_N}, \dots, \bar{w}_{Nn_N}^{Nn_N}\}$, and $\hat{\mathcal{W}} = \mathrm{diag}\{w_{11}^{11}, \dots, w_{11}^{Nn_N}, w_{12}^{11}, \dots, w_{12}^{Nn_N}, \dots, w_{Nn_N}^{Nn_N}\}$.

**Remark 1.** In Algorithm 1, (2a) is designed to meet the hard balance constraint under general directed topologies. It is straightforward to obtain from (2a) that $\sum_{ij\in\mathcal{V}_i} x_{ij}(k) = \sum_{ij\in\mathcal{V}_i} x_{ij}(0) = R_i$, $\forall k \in \mathbb{N}$ under arbitrary topologies. This can also be derived from the collective form of (2a), i.e., (4a), by noticing that $\mathbf{1}_{n_i}^{\mathrm{T}} \mathcal{L}_i^{\mathrm{o}} = 0$, which is the key property for making the hard balance constraint satisfied under (2a). Other matrices having this property can replace $\mathcal{L}_i^{\mathrm{o}}$ in (4a) to achieve $\sum_{ij\in\mathcal{V}_i} x_{ij}(k) = \sum_{ij\in\mathcal{V}_i} x_{ij}(0)$, $\forall k \in \mathbb{N}$ under general directed topologies. They might however, make the algorithm unable to comply with the demand of distributed computation. Therefore, the out-Laplacian matrix is introduced to ensure the distributed design of (2a), and meanwhile deal with the hard balance constraint and general directed topologies.

**Remark 2.** Observing the collective form (4a), the out-Laplacian matrix $\mathcal{L}_i^{\mathrm{o}}$ is in fact a singular linear transformation that builds the relation between $x_{ij}(0) - x_{ij}(k)$ and the virtual variable $z_{ij}(k)$. Under (2a), the original problem can be transformed into an equivalent problem regarding the decision variable $\boldsymbol{z}_i$ and the composite objective function $f_i(\boldsymbol{x}_1(0) - \mathcal{L}_1^{\mathrm{o}} \boldsymbol{z}_1, \dots, \boldsymbol{x}_N(0) - \mathcal{L}_N^{\mathrm{o}} \boldsymbol{z}_N)$ of each group $i$. To use the pseudo-gradient descent method to design the iteration of $\boldsymbol{z}_i$, the information of $\mathcal{L}_i^{\mathrm{oT}} \frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x})$ is required. Following this strategy, the value of $\sum_{im\in\mathcal{N}_{ij}^{\mathrm{io}}} (\frac{\partial f_i}{\partial x_{ij}}(\boldsymbol{x}) - \frac{\partial f_i}{\partial x_{im}}(\boldsymbol{x}))$ is desired for agent $ij$. However, since it is related to the group-level objective function, this information is unavailable to agent $ij$. To overcome this issue, (2c) and (2d) propose leader-follower consensus tracking protocols, where the auxiliary variables $s_{ij}^{pq}$ and $\boldsymbol{\eta}_{ij}^{il}$ are calculated by agent $ij$, and are estimates of $x_{pq}$ and

$\frac{\partial f_{il}}{\partial \boldsymbol{x}_i}(\boldsymbol{x})$, respectively; and therefore, the last term in (2b) can be viewed as an estimate of $\alpha\sum_{im\in\mathcal{N}_{ij}^{\text{io}}}(\frac{\partial f_i}{\partial x_{ij}}(\boldsymbol{x})-\frac{\partial f_i}{\partial x_{im}}(\boldsymbol{x}))$. In the proposed algorithm, (2a) and (2b) require each agent to know who are its intra-group out-neighbors and how many intra-group out-neighbors it has. This appears to be a limitation, albeit a modest one.

### 3.2. Convergence analysis

In this subsection, convergence analysis of Algorithm 1 is presented.

From (4a) and (4b), we have

$$\begin{aligned}\boldsymbol{x}_i(k+1)-\boldsymbol{x}_i(k) &=-\alpha\mathcal{L}_i^{\text{o}}\breve{\mathcal{L}}_i^{\text{o}}(I_{n_i}\otimes(\mathbf{1}_{n_i}^T\otimes I_{n_i}))\boldsymbol{\eta}_i(k),\\ \boldsymbol{x}(k+1)-\boldsymbol{x}(k) &=-\alpha\hat{\mathcal{L}}^{\text{o}}\hat{\breve{\mathcal{L}}}^{\text{o}}\hat{I}\boldsymbol{\eta}(k),\end{aligned}\tag{5}$$

where $\hat{\mathcal{L}}^{\text{o}}=\text{diag}\{\mathcal{L}_1^{\text{o}},\ldots,\mathcal{L}_N^{\text{o}}\}$, $\hat{\breve{\mathcal{L}}}^{\text{o}}=\text{diag}\{\breve{\mathcal{L}}_1^{\text{o}},\ldots,\breve{\mathcal{L}}_N^{\text{o}}\}$, and $\hat{I}=\text{diag}\{(I_{n_i}\otimes(\mathbf{1}_{n_i}^T\otimes I_{n_i}))\}$.

Define the estimate error $\boldsymbol{e}_s(k)=\boldsymbol{s}(k)-\mathbf{1}_{n_{\text{sum}}}\otimes\boldsymbol{x}(k)$. Then, one can derive from (5) and (4d) that

$$\begin{aligned}&\boldsymbol{e}_s(k+1)\\ &=(\mathcal{W}\otimes I_{n_{\text{sum}}}+\bar{\mathcal{W}})(\boldsymbol{s}(k)-\mathbf{1}_{n_{\text{sum}}}\otimes\boldsymbol{x}(k))\\ &\quad+(\mathcal{W}\otimes I_{n_{\text{sum}}}+\bar{\mathcal{W}}+\hat{\mathcal{W}})\big(\mathbf{1}_{n_{\text{sum}}}\otimes\boldsymbol{x}(k)\big)-\mathbf{1}_{n_{\text{sum}}}\otimes\boldsymbol{x}(k+1)\\ &=\mathcal{H}_{\mathcal{W}}\boldsymbol{e}_s(k)+\mathbf{1}_{n_{\text{sum}}}\otimes(\alpha\hat{\mathcal{L}}^{\text{o}}\hat{\breve{\mathcal{L}}}^{\text{o}}\hat{I}\boldsymbol{\eta}(k)),\end{aligned}\tag{6}$$

where $\mathcal{H}_{\mathcal{W}}=\mathcal{W}\otimes I_{n_{\text{sum}}}+\bar{\mathcal{W}}$, and the last equality is obtained by noticing that $(\mathcal{W}\otimes I_{n_{\text{sum}}}+\bar{\mathcal{W}}+\hat{\mathcal{W}})(\mathbf{1}_{n_{\text{sum}}}\otimes\boldsymbol{x})=\mathbf{1}_{n_{\text{sum}}}\otimes\boldsymbol{x}$. Under Assumption 1, it is not difficult to obtain from Gershgorin's Circle Theorem that $\mathcal{H}_{\mathcal{W}}$ is a Schur matrix. Thus, there exist a symmetric positive definite matrix $\mathcal{X}_{\mathcal{W}}$ such that $\mathcal{H}_{\mathcal{W}}^{\text{T}}\mathcal{X}_{\mathcal{W}}\mathcal{H}_{\mathcal{W}}-\mathcal{X}_{\mathcal{W}}=-I_{n_{\text{sum}}^2}$ (Chen, 1999).

Define the estimate error $\boldsymbol{e}_{\eta_i}(k)=\boldsymbol{\eta}_i(k)-\mathbf{1}_{n_i}\otimes\mathbf{Q}_i(\mathbf{1}_{n_i}\otimes\boldsymbol{x}(k))$. It follows from (4c) that

$$\begin{aligned}&\boldsymbol{e}_{\eta_i}(k+1)\\ &=(\mathcal{H}_{\mathcal{R}_i}\otimes I_{n_i})\boldsymbol{e}_{\eta_i}(k)+(\hat{\mathcal{R}}_i\otimes I_{n_i}) &(7)\\ &\quad\times\Big(\mathbf{1}_{n_i}\otimes\big(\mathbf{Q}_i(\boldsymbol{s}_i(k))-\mathbf{Q}_i(\mathbf{1}_{n_i}\otimes\boldsymbol{x}(k))\big)\Big) &(8)\\ &\quad-\mathbf{1}_{n_i}\otimes\big(\mathbf{Q}_i(\mathbf{1}_{n_i}\otimes\boldsymbol{x}(k+1))-\mathbf{Q}_i(\mathbf{1}_{n_i}\otimes\boldsymbol{x}(k))\big), &(9)\end{aligned}$$

where $\mathcal{H}_{\mathcal{R}_i}=\mathcal{R}_i\otimes I_{n_i}+\bar{\mathcal{R}}_i$, and the equality is obtained from $\big((\mathcal{R}_i\otimes I_{n_i}+\bar{\mathcal{R}}_i+\hat{\mathcal{R}}_i)\otimes I_{n_i}\big)\big(\mathbf{1}_{n_i}\otimes\mathbf{Q}_i(\mathbf{1}_{n_i}\otimes\boldsymbol{x})\big)=\mathbf{1}_{n_i}\otimes\mathbf{Q}_i(\mathbf{1}_{n_i}\otimes\boldsymbol{x})$. Similar to $\mathcal{H}_{\mathcal{W}}$, $\mathcal{H}_{\mathcal{R}_i}$ is also a Schur matrix under Assumption 1, and there exist a symmetric positive definite matrix $\mathcal{X}_{\mathcal{R}_i}$ such that $\mathcal{H}_{\mathcal{R}_i}^{\text{T}}\mathcal{X}_{\mathcal{R}i}\mathcal{H}_{\mathcal{R}_i}-\mathcal{X}_{\mathcal{R}i}=-I_{n_i^2}$.

From the above error systems, one can establish the convergence of the algorithm with sufficiently small $\alpha$. The result is stated as follows, with the proof reported in Appendix C.

**Theorem 1.** *Suppose that Assumptions 1–3 hold. The collective state $\boldsymbol{x}$ converges linearly to the NE of RAG (1) under Algorithm 1, if $\alpha$ satisfies*

$$\alpha\leq\min\left\{\frac{\gamma_\eta}{8\mu\max_{i\in\mathcal{I}}\{n_i\|\mathcal{L}_i^{\text{o}}\breve{\mathcal{L}}_i^{\text{o}}\|^2\}},\frac{\mu}{2\sum_{i=1}^N l_i^2\|\mathcal{L}_i^{\text{o}}\|^2+\gamma_\eta\sum_{i=1}^N\Big(b_i\sum_{j=1}^{n_i}l_{ij}^2\Big)+\gamma_s n_{sum}b}\right\},\tag{10}$$

*where*

$$b=2\|\mathcal{H}_{\mathcal{W}}^T\mathcal{X}_{\mathcal{W}}\|^2+\|\mathcal{X}_{\mathcal{W}}\|,\ b_i=2n_i\big(2\left\|\mathcal{H}_{\mathcal{R}_i}^T\mathcal{X}_{\mathcal{R}i}\right\|^2+\|\mathcal{X}_{\mathcal{R}i}\|\big),$$

$$\gamma_\eta=4\max_{i\in\mathcal{I}}\left\{n_i\|\breve{\mathcal{L}}_i^{\text{o}}\|^2\right\},\ \gamma_s=4\gamma_\eta\max_{i\in\mathcal{I}}\left\{b_i\|\hat{\mathcal{R}}_i\|^2\max_{ij\in\mathcal{V}_i}\{l_{ij}^2\}\right\}.$$

*Moreover, the equality constraint in (1) is satisfied at every iteration.*

**Remark 3.** Theorem 1 shows that Algorithm 1 can achieve the linear convergence of the collective state to the exact NE of RAG (1), while also ensuring satisfaction of the hard balance constraint under general directed typologies. In RAG (1), the individual benefit of each agent may be influenced by the states of all the agents, including those within the same group. It is worth noting that there are scenarios where the individual benefit of each agent is independent of the states of other intra-group members. Recalling Example 1 in Section 2.2, the business lines within the company may be of different types, and the revenue of each business line is influenced solely by its own computing power and that of similar business lines from other companies. In such cases, problem (1) degenerates to the following form.

$$\begin{aligned}&\min_{\boldsymbol{x}_i}f_i(\boldsymbol{x})=\min_{\boldsymbol{x}_i}\sum_{il\in\mathcal{V}_i}f_{il}(x_{il},\boldsymbol{x}_{-i}),\ \forall i\in\mathcal{I},\\ &\text{s.t.}\ \sum_{ij\in\mathcal{V}_i}x_{ij}=R_i.\end{aligned}\tag{11}$$

Model (11) can be referred to as intra-independent resource allocation game (IIRAG), while model (1) can be referred to as intra-dependent resource allocation game (IDRAG). Obviously, IIRAG is a special case of IDRAG. In IIRAG (11), the following property holds:

$$\frac{\partial f_i}{\partial x_{ij}}(\boldsymbol{x})=\sum_{il\in\mathcal{V}_i}\frac{\partial f_{il}}{\partial x_{ij}}(x_{il},\boldsymbol{x}_{-i})=\frac{\partial f_{ij}}{\partial x_{ij}}(x_{ij},\boldsymbol{x}_{-i}),\tag{12}$$

allowing for a slight modification of the updating law (2) to save computational complexity:

$$x_{ij}(k+1)=x_{ij}(0)-d_{ij}^{\text{io}}z_{ij}(k+1)+\sum_{im\in\mathcal{N}_{ij}^{\text{i}}}z_{im}(k+1),\tag{13a}$$

$$z_{ij}(k+1)=z_{ij}(k)+\alpha\sum_{im\in\mathcal{N}_{ij}^{\text{io}}}\left(\psi_{ij}^{ij}(k)-\psi_{ij}^{im}(k)\right),\tag{13b}$$

$$\psi_{ij}^{il}(k+1)=\bar{r}_{ij}^{il}\psi_{ij}^{il}(k)+\sum_{im\in\mathcal{N}_{ij}^{\text{i}}}r_{ij}^{im}\psi_{im}^{il}(k)+r_{ij}^{il}\frac{\partial f_{il}}{\partial x_{il}}(s_{il}^{il}(k),\boldsymbol{s}_{il}^{-i}(k)),\quad\forall il\in\mathcal{V}_i,\tag{13c}$$

$$s_{ij}^{pq}(k+1)=\bar{w}_{ij}^{pq}s_{ij}^{pq}(k)+\sum_{lm\in\mathcal{N}_{ij}}w_{ij}^{lm}s_{lm}^{pq}(k)+w_{ij}^{pq}x_{pq}(k),\qquad\forall pq\in\mathcal{V},\tag{13d}$$

where $\boldsymbol{s}_{il}^{-i}=\text{col}(\boldsymbol{s}_{il}^1,\ldots,\boldsymbol{s}_{il}^{i-1},\boldsymbol{s}_{il}^{i+1},\ldots,\boldsymbol{s}_{il}^N)$, $\psi_{ij}^{il}$ is a scalar auxiliary variable, and other parameters are the same as in (2). Specifically, $\psi_{ij}^{il}$ is calculated by agent $ij$ to estimate $\frac{\partial f_{il}}{\partial x_{il}}(x_{il}^{il},\boldsymbol{x}_{-i})$ (or equivalently $\frac{\partial f_i}{\partial x_{il}}(\boldsymbol{x})$ in light of (12)), indicating that the last term in (13b) is an estimate of $\alpha\sum_{im\in\mathcal{N}_{ij}^{\text{io}}}(\frac{\partial f_i}{\partial x_{ij}}(\boldsymbol{x})-\frac{\partial f_i}{\partial x_{im}}(\boldsymbol{x}))$. The adjustment from (2) to (13) results in a reduction in the dimensionality of the auxiliary variables from $(1+n_i^2+n_{\text{sum}})$ to $(1+n_i+n_{\text{sum}})$.

**Corollary 1.** *Suppose that Assumptions 1–3 hold. Under the updating law (13), the collective state $\boldsymbol{x}$ converges linearly to the NE of IIRAG (11), if $\alpha$ satisfies*

$$\alpha\leq\min\left\{\frac{\gamma_\psi}{8\mu\max_{i\in\mathcal{I}}\left\{\left\|\mathcal{L}_i^{\text{o}}\breve{\mathcal{L}}_i^{\text{o}}\right\|^2\right\}},\frac{\mu}{2\sum_{i=1}^N l_i^2\|\mathcal{L}_i^{\text{o}}\|^2+\gamma_\psi\sum_{i=1}^N b_il_i^2+\gamma_{s_2}n_{sum}b}\right\},\tag{14}$$

*where $b, b_i$ are defined in Theorem 1, $\gamma_\psi=4\max_{i\in\mathcal{I}}\left\{\|\breve{\mathcal{L}}_i^{\text{o}}\|^2\right\}$, $\gamma_{s_2}=4\gamma_\psi\max_{i\in\mathcal{I}}\left\{b_il_i^2\|\hat{\mathcal{R}}_i\|^2\right\}$. Moreover, the equality constraint in (11) is satisfied at every iteration.*

The proof is similar to that of Theorem 1, and therefore omitted.

**Remark 4.** The proposed DRA algorithms are developed from a game-theoretic perspective, successfully addressing conflicts of interest among groups. Most existing works on DRA are concerned with only a single group of cooperative agents. In fact, the basic form of these models can be viewed as a special case of the IIRAG with $N=1$. In this sense, the proposed method generalizes the existing works on DRA from one single group to multiple interacting groups. Another distinct feature of the proposed method is that it can make the balance constraint inside each group satisfied during the whole process of NE seeking. This benefits from the proposed out-Laplacian framework, which is different from the primal–dual method (Bazaraa, Sherali, & Shetty, 2006). It is worth noting that, although the primal–dual method is an effective tool to deal with constraints, it is inadequate to address the demand of maintaining the balance at every iteration.

**Remark 5.** The RAG problem with hard balance constraint has been explored in Zhou et al. (2024), where distributed NE computation algorithms are proposed to achieve linear convergence of the collective state to the NE. However, these algorithms are limited to undirected graphs because they rely on the symmetry of Laplacian matrices of the group-level subgraphs. Consequently, they cannot be applicable to directed graphs, even those that are balanced. In contrast, distributed NE computation algorithms for RAG under balanced graphs are devised in Deng and Liu (2023). Nevertheless, these algorithms require the Laplacian matrices of the group-level subgraphs to have a left eigenvector of $\mathbf{1}$ associated with a zero eigenvalue, rendering them inapplicable to unbalanced graphs. Moreover, the hard balance constraint is not addressed in Deng and Liu (2023). The combination of unbalanced directed topologies and the hard balance constraint significantly complicates the design of distributed NE computation algorithms for RAG. Fortunately, the potential of the out-Laplacian matrix $\mathcal{L}_i^{\mathrm{o}}$ has been discovered. This leads to the formulation of the updating law (2), which effectively addresses the hard balance constraint and accommodates general directed topologies.

## 4. Reduced-order DRA algorithm for RAG

In the previous section, we introduced a new DRA algorithm, Algorithm 1, tailored for RAG (1) with the hard balance constraint and directed topologies. Notably, the dimensionality of auxiliary variables in Algorithm 1 is $(1+n_i^2+n_{\mathrm{sum}})$, leading to considerable computational complexity. To address this challenge, in this section, we develop a reduced-order DRA algorithm for RAG (1), with linear convergence also established.

### 4.1. Algorithm design

Instead of using the leader-follower protocol to estimate the partial derivatives of individual objective functions, a gradient-tracking law (15c) is developed, where $\xi_{ij}^{il}$ plays the role of tracking $\frac{v_{ij}}{n_i}\cdot\frac{\partial f_i}{\partial x_{il}}(\boldsymbol{x})$ with $v_{ij}$ being a positive constant to be introduced later. The design of (15b) still follows the out-Laplacian matrix based methodology illustrated before:

$$x_{ij}(k{+}1)=x_{ij}(0)-d_{ij}^{\mathrm{io}}z_{ij}(k{+}1)+\sum_{im\in\mathcal{N}_{ij}^{i}}z_{im}(k{+}1), \tag{15a}$$

$$z_{ij}(k{+}1)=z_{ij}(k)+\beta\sum_{im\in\mathcal{N}_{ij}^{\mathrm{io}}}\left(\xi_{ij}^{ij}(k)-\xi_{ij}^{im}(k)\right), \tag{15b}$$

$$\xi_{ij}^{il}(k{+}1)=\sum_{im\in\mathcal{N}_{ij}^{i}}c_{ij}^{im}\xi_{im}^{il}(k)+\frac{\partial f_{ij}}{\partial x_{il}}\big(\boldsymbol{s}_{ij}(k{+}1)\big)-\frac{\partial f_{ij}}{\partial x_{il}}\big(\boldsymbol{s}_{ij}(k)\big),\quad \forall il\in\mathcal{V}_i, \tag{15c}$$

$$s_{ij}^{pq}(k{+}1)=\bar{w}_{ij}^{pq}s_{ij}^{pq}(k)+\sum_{lm\in\mathcal{N}_{ij}}w_{ij}^{lm}s_{lm}^{pq}(k)+w_{ij}^{pq}x_{pq}(k),\qquad \forall pq\in\mathcal{V}. \tag{15d}$$

---

**Algorithm 2** Reduced-order DRA algorithm for the RAG

---

For agent $ij\in\mathcal{V}$

1: **Input:** intra-group out-degree $d_{ij}^{\mathrm{io}}$, weights $c_{il}^{ij}$, $\forall il\in\mathcal{V}_i$ and $w_{ij}^{pq}$, $\forall pq\in\mathcal{V}$ defined in Section 2.1, weights $\bar{w}_{ij}^{pq}$, $\forall pq\in\mathcal{V}$ defined in (3), parameter $\beta$.
2: **Initialize:** $x_{ij}(0)=R_{ij}$, $z_{ij}(0)=0$, $\xi_{ij}^{il}(0)=\frac{\partial f_{ij}}{\partial x_{il}}(\boldsymbol{s}_{ij}(0))$, $\forall il\in\mathcal{V}_i$, and arbitrary $\boldsymbol{s}_{ij}(0)\in\mathbb{R}^{n_{\mathrm{sum}}}$.
3: **for** $k=0,1,2,\ldots$ **do**
4: Send $x_{ij}(k)$, $\boldsymbol{s}_{ij}(k)$ to its out-neighbor $uv$ for all $uv\in\mathcal{N}_{ij}^{\mathrm{o}}$, and send $z_{ij}(k)$, $c_{ih}^{ij}\xi_{ij}^{i1}(k),\ldots,c_{ih}^{ij}\xi_{ij}^{in_i}(k)$ to its intra-group out-neighbor $ih$ for all $ih\in\mathcal{N}_{ij}^{\mathrm{io}}$. Receive $x_{pq}(k)$, $\boldsymbol{s}_{pq}(k)$ from its in-neighbor $pq$ for all $pq\in\mathcal{N}_{ij}$, and receive $z_{im}(k)$, $c_{ij}^{im}\xi_{im}^{i1}(k),\ldots,c_{ij}^{im}\xi_{im}^{in_i}(k)$ from its intra-group in-neighbor $im$ for all $im\in\mathcal{N}_{ij}^{i}$.
5: Compute $z_{ij}(k+1)$ according to (15b).
6: **for** $pq\in\mathcal{V}$ **do**
7: Compute $s_{ij}^{pq}(k+1)$ according to (15d).
8: **end for**
9: **for** $il\in\mathcal{V}_i$ **do**
10: Compute $\xi_{ij}^{il}(k+1)$ according to (15c).
11: **end for**
12: Compute $x_{ij}(k+1)$ according to (15a).
13: **end for**

---

The collective form of (15) is as follows $\forall i\in\mathcal{I}$:

$$\boldsymbol{x}_i(k{+}1)=\boldsymbol{x}_i(0)-\mathcal{L}_i^{\mathrm{o}}\boldsymbol{z}_i(k{+}1), \tag{16a}$$

$$\boldsymbol{z}_i(k{+}1)=\boldsymbol{z}_i(k)+\beta\check{\mathcal{L}}_i^{\mathrm{o}}\boldsymbol{\xi}_i(k), \tag{16b}$$

$$\boldsymbol{\xi}_i(k{+}1)=(\mathcal{C}_i\otimes I_{n_i})\boldsymbol{\xi}_i(k)+\mathbf{Q}_i(\boldsymbol{s}_i(k{+}1))-\mathbf{Q}_i(\boldsymbol{s}_i(k)), \tag{16c}$$

$$\boldsymbol{s}(k{+}1)=(\mathcal{W}\otimes I_{n_{\mathrm{sum}}}+\bar{\mathcal{W}})\boldsymbol{s}(k)+\hat{\mathcal{W}}\big(\mathbf{1}_{n_{\mathrm{sum}}}\otimes\boldsymbol{x}(k)\big), \tag{16d}$$

where $\boldsymbol{\xi}_i=\mathrm{col}(\xi_{i1}^{i1},\xi_{i1}^{i2},\ldots,\xi_{i1}^{in_i},\xi_{i2}^{i1},\ldots,\xi_{in_i}^{in_i})\in\mathbb{R}^{n_i^2}$, $\beta$ is a small constant step size to be determined later, and $\mathcal{C}_i$ is the column-stochastic matrix defined in Section 2.1. Other vectors and matrices are also defined as before.

**Remark 6.** The difference between Algorithms 1 and 2 lies in the mechanism of estimating partial derivatives of objective functions. Algorithm 1 employs the leader-follower tracking mechanism, where $\boldsymbol{\eta}_{ij}^{il}$ in (2c) acts as followers tracking partial derivatives of intra-group individual objective functions $\frac{\partial f_{il}}{\partial \boldsymbol{x}_i}(\boldsymbol{s}_{il})$; while Algorithm 2 adopts the gradient-tracking mechanism, where $\xi_{ij}^{il}$ in (15c) is designed to track weighted partial derivatives of the group-level objective function $\frac{v_{ij}}{n_i}\cdot\frac{\partial f_i}{\partial x_{il}}(\boldsymbol{s}_{ij})$. The design of (15c) is similar to the gradient tracking method in distributed optimization (Nedić, Olshevsky, & Shi, 2017; Xin & Khan, 2020), where the estimation of the average gradient is constructed. Compared with the gradient-tracking mechanism, the leader-follower tracking mechanism is more straightforward and has the advantage of less complicated convergence analysis. Moreover, the initial values of the estimators under the leader-follower tracking mechanism, known as $\boldsymbol{\eta}_{ij}^{il}(0)$ in Algorithm 1, can be arbitrarily chosen; while those under the gradient-tracking mechanism, known as $\xi_{ij}^{il}(0)$, should be particularly set as $\frac{\partial f_{ij}}{\partial x_{il}}(\boldsymbol{s}_{ij}(0))$. However, the disadvantage of the leader-follower tracking mechanism is the heavier computational complexity in the RAG (1) compared with the gradient-tracking mechanism. The total dimensionality of the updating variables decreases from $(2+n_i^2+n_{\text{sum}})$ in Algorithm 1 to $(2+n_i+n_{\text{sum}})$ in Algorithm 2. Besides, under the leader-follower tracking mechanism in Algorithm 1, the agents need to share the gradient information through the network, potentially compromising the privacy of individual cost functions. In contrast, the gradient-tracking mechanism in Algorithm 2 eliminates the need to transmit the gradient information of individual cost functions, thus safeguarding their privacy. This appears to be another advantage of the gradient-tracking mechanism over the leader-follower tracking mechanism. Modifying Algorithm 1 to incorporate privacy-preserving properties is an intriguing challenge that we plan to explore in our future work.

**Remark 7.** Note that in Algorithm 2, all iterative variables ($x_{ij}$, $\boldsymbol{s}_{ij}$, $z_{ij}$, $c_{ih}^{ij}\xi_{il}^{ij}$) need to be transmitted over the network. Such extensive communication among agents imposes a significant communication burden. However, in the current algorithmic design, the transmission of these variables is necessary. Specifically, the transmission of variables $\boldsymbol{s}_{ij}$ and $x_{pq}$ is essential for realizing distributed estimation of $\boldsymbol{x}$ to compute gradient information locally; the transmission of variables $c_{ih}^{ij}\xi_{il}^{ij}$ is necessary for estimating the group-level gradient $\frac{v_{ij}}{n_i}\cdot\frac{\partial f_i}{\partial x_{il}}(\boldsymbol{x})$ for performing the gradient descent algorithm in (13b); and the transmission of variables $z_{ij}$ is crucial for ensuring hard balance constraint during the iteration process. While reducing the communication burden is a significant concern, introducing event-triggering mechanisms appears to be a potential solution. In the future, we plan to design event-triggered DRA algorithms for the RAG problem.

### 4.2. Convergence analysis

For notational convenience, define

$$\bar{\boldsymbol{\xi}}_i=\frac{1}{n_i}\left(\mathbf{1}_{n_i}^T\otimes I_{n_i}\right)\boldsymbol{\xi}_i\in\mathbb{R}^{n_i},$$
$$\bar{\mathbf{Q}}_i(\cdot)=\frac{1}{n_i}(\mathbf{1}_{n_i}^T\otimes I_{n_i})\mathbf{Q}_i(\cdot)\in\mathbb{R}^{n_i}. \tag{17}$$

From the initialization in Algorithm 2, one has $\boldsymbol{\xi}_i(0)=\mathbf{Q}_i(\boldsymbol{s}_i(0))$. Then, one can derive from (16c) and $\mathbf{1}_{n_i}^T\mathcal{C}_i=\mathbf{1}_{n_i}^T$ that

$$\bar{\boldsymbol{\xi}}_i(k)=\bar{\mathbf{Q}}_i(\boldsymbol{s}_i(k)),\quad \forall k\in\mathbb{N}. \tag{18}$$

Define the estimate errors $\boldsymbol{e}_{\xi_i}(k)=\boldsymbol{\xi}_i(k)-\boldsymbol{v}_i\otimes\bar{\boldsymbol{\xi}}_i(k)$, where $\boldsymbol{v}_i=\text{col}(v_{i1},\ldots,v_{in_i})$ with $\mathbf{1}_{n_i}^T\boldsymbol{v}_{n_i}=n_i$ and $v_{ij}>0, \forall j\in\mathcal{V}_i$ is the right positive eigenvector of $\mathcal{C}_i$ corresponding to the eigenvalue 1, i.e., $\mathcal{C}_i\boldsymbol{v}_i=\boldsymbol{v}_i$. The existence of $\boldsymbol{v}_i$ is guaranteed by Assumption 1 that $\mathcal{G}_i$ is strongly connected (Mei, Ren, & Chen, 2016). From (17),

$$\boldsymbol{e}_{\xi_i}(k)=(\bar{I}_i\otimes I_{n_i})\boldsymbol{\xi}_i(k), \tag{19}$$

where $\bar{I}_i=I_{n_i}-\frac{\boldsymbol{v}_i\mathbf{1}_{n_i}^T}{n_i}$. Define $\bar{\mathcal{C}}_i=\mathcal{C}_i-\frac{\boldsymbol{v}_i\mathbf{1}_{n_i}^T}{n_i}$. By noting (16c), (19), and $\bar{I}_i\mathcal{C}_i=\bar{\mathcal{C}}_i$, one can derive

$$\begin{aligned}\boldsymbol{e}_{\xi_i}(k+1)=&\left(\bar{\mathcal{C}}_i\otimes I_{n_i}\right)\boldsymbol{e}_{\xi_i}(k)\\&+\left(\bar{I}_i\otimes I_{n_i}\right)\left(\mathbf{Q}_i(\boldsymbol{s}_i(k+1))-\mathbf{Q}_i(\boldsymbol{s}_i(k))\right).\end{aligned} \tag{20}$$

Since $\lim_{k\to\infty}(\mathcal{C}_i)^k=\boldsymbol{v}_i\mathbf{1}_{n_i}^T/n_i$ under Assumption 1 (Ren & Beard, 2005), one has $\lim_{k\to\infty}(\bar{\mathcal{C}}_i)^k=0$, implying that $\bar{\mathcal{C}}_i$ is a Schur matrix. Therefore, there exists a symmetric positive definite matrix $\mathcal{X}_{\mathcal{C}i}$ such that $\bar{\mathcal{C}}_i^T\mathcal{X}_{\mathcal{C}i}\bar{\mathcal{C}}_i-\mathcal{X}_{\mathcal{C}i}=-I_{n_i}$ (Chen, 1999).

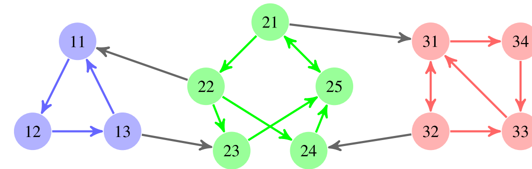



**Fig. 1.** The communication topology among agents.

**Theorem 2.** *Suppose that Assumptions 1–3 hold. The collective state $\boldsymbol{x}$ converges linearly to the NE of the RAG (1) under Algorithm 2, if $\beta$ satisfies*

$$\beta\leq\min\left\{\frac{\gamma_\xi}{8\mu\max_{i\in\mathcal{I}}\{\|\mathcal{L}_i^{\mathrm{o}}\breve{\mathcal{L}}_i^{\mathrm{o}}\|^2\}},\frac{\mu}{2\sum_{i=1}^N\|\hat{\boldsymbol{v}}_i\|\left\|\mathcal{L}_i^{\mathrm{o}}\right\|^2 l_i^2/n_i+2\tilde{\gamma}_s n_{sum}b},\frac{\tilde{\gamma}_s}{8\mu\max_{i\in\mathcal{I}}\left\{\|\mathcal{L}_i^{\mathrm{o}}\hat{\boldsymbol{v}}_i\mathcal{L}_i^{\mathrm{o}T}\|^2\sum_{j=1}^{n_i}l_{ij}^2/n_i^2\right\}}\right\}, \tag{21}$$

*where $\hat{\boldsymbol{v}}_i=\text{diag}\{v_{i1},\ldots,v_{in_i}\}$, $\gamma_\xi=4\max_{i\in\mathcal{I}}\{n_i\|\hat{\boldsymbol{v}}_i^{-1}\|\|\breve{\mathcal{L}}_i^{\mathrm{o}}\|^2\}$, $\tilde{\gamma}_s=4\Big(\max_{i\in\mathcal{I}}\Big\{\|\hat{\boldsymbol{v}}_i\|\left\|\mathcal{L}_i^{\mathrm{o}}\right\|^2\sum_{j=1}^{n_i}l_{ij}^2/n_i\Big\}+\gamma_\xi\tilde{b}\|I_{n_{\text{sum}}^2}-\mathcal{H}_{\mathcal{W}}\|^2\Big)$, $\tilde{b}=\max_{i\in\mathcal{I}}\Big\{\big(2\|\bar{\mathcal{C}}_i^T\mathcal{X}_{\mathcal{C}i}\bar{I}_i\|^2+\|\bar{I}_i^T\mathcal{X}_{\mathcal{C}i}\bar{I}_i\|\big)\max_{ij\in\mathcal{V}_i}\{l_{ij}^2\}\Big\}$, and $b$ is defined in Theorem 1. Moreover, the equality constraint in (1) is satisfied at every iteration.*

The proof of Theorem 2 is reported in Appendix D.

## 5. Numerical examples

In this section, we examine the performance of the proposed DRA algorithms in a networked Cournot competition game. The game is a variant of the one discussed in Pavel (2020). It involves three internet companies (groups), each with a different number of business lines (agents). We denote these companies as $\mathcal{I}=\{1,2,3\}$ and their corresponding business lines as $\mathcal{V}_1=\{11,12,13\}$, $\mathcal{V}_2=\{21,22,23,24,25\}$, $\mathcal{V}_3=\{31,32,33,34\}$, which aligns with the definitions provided in Section 2. The business lines strategically disseminate certain information to selected entities while withholding it from others to safeguard strategic interests and maintain privacy. Additionally, they employ connectionless communication protocols that allow for unidirectional data transmission, which are more efficient than two-way protocols with extensive handshakes in bandwidth-limited environments. These factors result in the communication topology among the agents being a directed graph, as illustrated in Fig. 1.

The companies compete in multiple sub-markets by adjusting the allocation of computing power among their business lines. The computing power possessed by the three companies are $R_1=100$, $R_2=200$, and $R_3=180$, respectively. Let $\mathcal{M}_l$ represent the set of business lines serving in the $l$th sub-market, and $h_{ij}$ represent the index of the sub-market that business line $ij$ served in, i.e., $ij\in\mathcal{M}_{h_{ij}}$. We use $x_{ij}$ to denote the amount of computing power allocated to business line $ij$, and $k_{ij}$ to denote the amount of computing power consumed per unit of service in business line $ij$. The price per unit of service in the $l$th sub-market depends on the total volume of services provided by all business lines in that sub-market, calculated by $p^l=600-\sum_{pq\in\mathcal{M}_l}x_{pq}/k_{pq}$. The revenue of business line $ij\in\mathcal{M}_l$ can be calculated by $\mathbf{r}_{ij}=p^l x_{ij}/k_{ij}$. Additionally, the cost of computing power consumed by business line $ij$ is given by $\mathbf{c}_{ij}=5x_{ij}^2-10x_{ij}+20$. The payoff of business line

$ij$ is determined by the difference between its revenue and cost, which can be expressed as $\mathbf{r}_{ij} - \mathbf{c}_{ij}$. Each company aims to maximize the sum of payoffs across all its business lines. Therefore, in the minimization setting, the local objective function of business line $ij$ can be written as $5x_{ij}^2 - 10x_{ij} + 20 - (600 - \sum_{pq\in\mathcal{M}_{h_{ij}}} \frac{x_{pq}}{k_{pq}})\frac{x_{ij}}{k_{ij}}$. Then, the game is expressed as:

$$\min_{x_{ij}} \sum_{il\in\mathcal{V}_i} \left( 5x_{il}^2 - 10x_{il} + 20 - (600 - \sum_{pq\in\mathcal{M}_{h_{il}}} \frac{x_{pq}}{k_{pq}})\frac{x_{il}}{k_{il}} \right), \quad \forall ij \in \mathcal{V},$$

$$s.t. \sum_{ij\in\mathcal{V}_i} x_{ij} = R_i, \quad \forall i \in \mathcal{I}.$$

Consider there are two sub-markets, denoted as $\mathcal{M}_1 = \{11, 12, 21, 22, 31, 32\}$, $\mathcal{M}_2 = \{13, 23, 24, 25, 33, 34\}$. Set $k_{11} = 1$, $k_{12} = 2$, $k_{13} = 3$, $k_{21} = 1$, $k_{22} = 2$, $k_{23} = 1$, $k_{24} = 2$, $k_{25} = 3$, $k_{31} = 1$, $k_{32} = 2$, $k_{33} = 1$, $k_{34} = 2$. It can be directly calculated in a centralized manner that the NE in this RAG is $\boldsymbol{x}^* = [45.71, 28.56, 25.73, 50.36, 33.57, 52.63, 34.71, 28.73, 52.16, 35.5, 55.28, 37.06]^{\mathrm{T}}$.

We initialize the collective state as $\boldsymbol{x}(0) = [20, 70, 10, 30, 40, 50, 20, 60, 30, 20, 50, 80]^T$, and set the parameters $w_{ij}^{pq}$, $r_{ij}^{il}$, $c_{ij}^{il}$ associated with the weighted adjacency matrices $\mathcal{W}, \mathcal{R}_i, \mathcal{C}_i$ as in Section 2.1 with $h_{ij} = 1.1$ and $h_{ij}^i = 1.1$. Then, we execute Algorithms 1 and 2 to evaluate their effectiveness.

The simulation results under Algorithm 1 with a step size of $\alpha = 0.0005$ and Algorithm 2 with a step size of $\beta = 0.01$ are depicted in Figs. 2 and 3, respectively. These figures demonstrate that under both algorithms, the collective state successfully converges to the NE. Additionally, the sum of agent states in each group consistently matches the available computing resources at every iteration, satisfying the hard balance constraint throughout the distributed NE computation process. Furthermore, the results under Algorithm 2 indicate a faster convergence rate compared to Algorithm 1.

Given the pivotal role of step sizes in influencing convergence rates, we will further discuss the convergence behavior of Algorithms 1 and 2 regarding the choice of step sizes $\alpha$ and $\beta$. From Theorem 1, it can be inferred that there exists a critical value $\alpha^*$ for the step size $\alpha$, ensuring convergence if $\alpha \in (0, \alpha^*)$. A step size that is too small may result in slow convergence, while excessively large values can lead to divergence. Initially, increasing the step size within the range $(0, \alpha^*)$ may enhance convergence. However, beyond a specific threshold within this range, further increments may cause overshooting or oscillations, thereby diminishing the convergence rate. Beyond this range, oscillations and divergence ensue. Notably, the expression on the right-hand side of inequality (10) does not denote the critical value; rather, it offers a sufficient but not necessary condition for convergence. Through trial, one can find that $\alpha^*$ for Algorithm 1 is approximately 0.00104, and setting $\alpha = 0.0005$ yields a good convergence rate, as depicted in Fig. 2. Similarly, for Algorithm 2, one can find that $\beta^*$ is approximately 0.0501, and setting $\beta = 0.01$ achieves satisfactory convergence. Comparing these two cases, it seems that Algorithm 2's broader step size range gives it an advantage over Algorithm 1 in achieving faster convergence rates when both use appropriate step sizes.

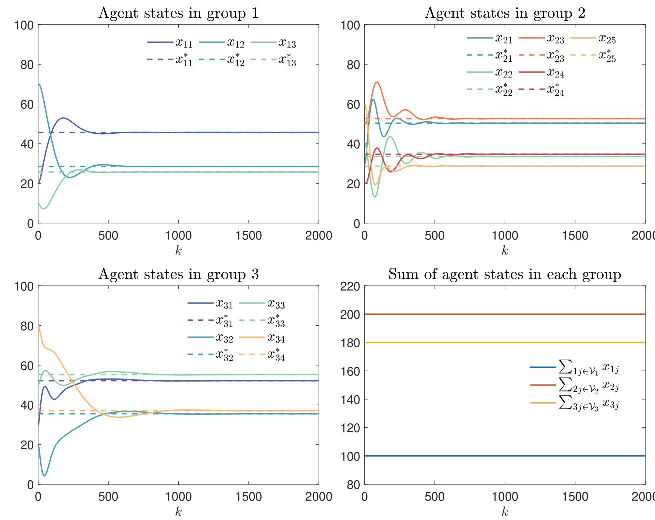



**Fig. 2.** Simulation results under Algorithm 1.

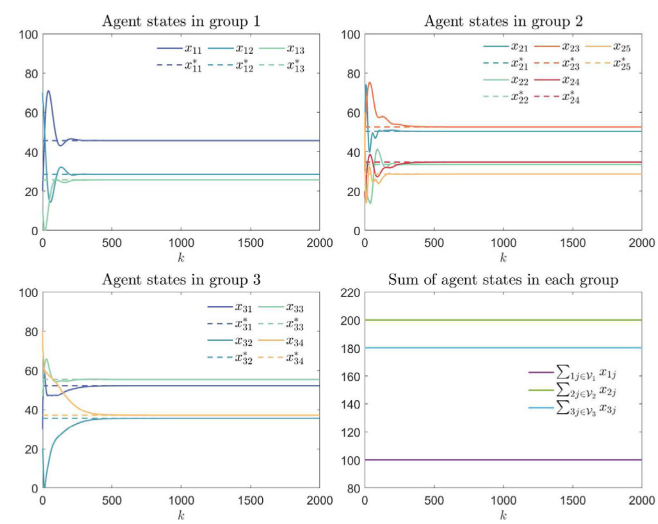



**Fig. 3.** Simulation results under Algorithm 2.

## 6. Conclusion

In this paper, a new out-Laplacian matrix based design approach was developed for DRA over multiple groups of agents with hard balance constraints under general directed topologies. Considering the cases that the individual benefit of each agent is explicitly influenced by the states of itself and agents in other groups, and by the states of all agents, two multi-group RAG models were studied, with three DRA algorithms developed. The proposed algorithms can be classified into two categories based on the mechanisms in estimating partial derivatives. One is the leader-follower consensus tracking mechanism, which enjoys convenience in convergence analysis, and the other is the gradient-tracking mechanism, which can reduce the computational complexity in IDRAG. Under the proposed algorithms, the hard balance constraint was satisfied, and linear convergence of the agent states to the exact NE was rigorously demonstrated. Future research will extend the proposed methods to accommodate additional general local convex constraints, with particular emphasis on preserving satisfaction of those constraints throughout the iteration process and broadening their applicability to diverse online resource allocation scenarios.

## Appendix A. Proof of Lemma 1

Under Assumption 2, one can get from Karush–Kuhn–Tucker optimality condition (Bazaraa et al., 2006) that, a vector $\boldsymbol{x}^* \in \Omega$ is the NE of RAG (1) if and only if there exist $\lambda_1, \ldots, \lambda_N \in \mathbb{R}$

such that $\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}^*)+\lambda_i \mathbf{1}_{n_i} = \mathbf{0}_{n_i}$, $\forall i \in \mathcal{I}$, which is equivalent to $\mathcal{L}_i^{o\mathrm{T}} \frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}^*) = \mathbf{0}_{n_i}$, $\forall i \in \mathcal{I}$ under Assumption 1.

## Appendix B. Two useful lemmas

**Lemma 2** (Horn, Rhee, & Wasin, 1998). *Denote the eigenvalues of a symmetric matrix $A\in\mathbb{R}^{n\times n}$ by $\lambda_1,\ldots,\lambda_n$, which satisfies $\lambda_1\leq\ldots\leq\lambda_n$. Let $\upsilon_j$ denote the eigenvector of $A$ associated with the eigenvalue $\lambda_j$. Then, the following inequality holds for $i=1,\ldots,n$: $y^TAy\geq\lambda_i y^Ty$, $\forall y\in\{y|y\perp\upsilon_j, j=1,\ldots,i-1\}$.*

**Lemma 3** (Young's Inequality). *Let $p, q$ be positive real numbers satisfying $\frac{1}{p}+\frac{1}{q}=1$. Then, for any $a\geq 0, b\geq 0$, $ab\leq\frac{a^p}{p}+\frac{b^q}{q}$.*

## Appendix C. Proof of Theorem 1

Define $V_s(k)=\boldsymbol{e}_s^{\mathrm{T}}(k)\mathcal{X}_{\mathcal{W}}\boldsymbol{e}_s(k)$, where $\mathcal{X}_{\mathcal{W}}$ has been given below (6). From (6), one can derive that

$$\begin{aligned}
&V_s(k+1)-V_s(k)\\
&\leq-\|\boldsymbol{e}_s(k)\|^2+2\sqrt{n_{\mathrm{sum}}}\|\mathcal{H}_{\mathcal{W}}^{\mathrm{T}}\mathcal{X}_{\mathcal{W}}\|\|\boldsymbol{e}_s(k)\|\|\alpha\hat{\mathcal{L}}^o\check{\mathcal{L}}^o\hat{I}\boldsymbol{\eta}(k)\|\\
&\quad+n_{\mathrm{sum}}\|\mathcal{X}_{\mathcal{W}}\|\left\|\alpha\hat{\mathcal{L}}^o\check{\mathcal{L}}^o\hat{I}\boldsymbol{\eta}(k)\right\|^2\\
&\leq-\frac{1}{2}\|\boldsymbol{e}_s(k)\|^2+n_{\mathrm{sum}}b\left\|\alpha\hat{\mathcal{L}}^o\check{\mathcal{L}}^o\hat{I}\boldsymbol{\eta}(k)\right\|^2,
\end{aligned}\tag{C.1}$$

where $b$ is defined in Theorem 1 .

Define $V_\eta(k)=\boldsymbol{e}_\eta^{\mathrm{T}}(k)\mathcal{X}_{\mathcal{R}}\boldsymbol{e}_\eta(k)$, where $\mathcal{X}_{\mathcal{R}}=\mathrm{diag}\{\mathcal{X}_{\mathcal{R}_1}\otimes I_{n_1},\ldots,\mathcal{X}_{\mathcal{R}_N}\otimes I_{n_N}\}$ with $\mathcal{X}_{\mathcal{R}_i}$ defined below (9). From (9), one can get

$$\begin{aligned}
&V_\eta(k+1)-V_\eta(k)\\
&\leq\sum_{i=1}^{N}\Bigg(-\frac{1}{2}\left\|\boldsymbol{e}_{\eta_i}(k)\right\|^2+\left(2\left\|\mathcal{H}_{\mathcal{R}_i}^T\mathcal{X}_{\mathcal{R}i}\right\|^2+\|\mathcal{X}_{\mathcal{R}i}\|\right)\\
&\times\Big\|(\hat{\mathcal{R}}_i\otimes I_{n_i})\big(\mathbf{1}_{n_i}\otimes(\mathbf{Q}_i(\boldsymbol{s}_i(k)))-\mathbf{Q}_i(\mathbf{1}_{n_i}\otimes\boldsymbol{x}(k))\big)\\
&-\mathbf{1}_{n_i}\otimes\Big(\mathbf{Q}_i(\mathbf{1}_{n_i}\otimes\boldsymbol{x}(k+1))-\mathbf{Q}_i(\mathbf{1}_{n_i}\otimes\boldsymbol{x}(k))\Big)\Big\|^2\Bigg),
\end{aligned}\tag{C.2}$$

where Young's inequality is used similarly as in (C.1).

Under Assumption 2, one has

$$\begin{aligned}
&\|\mathbf{Q}_i(\boldsymbol{s}_i(k))-\mathbf{Q}_i(\mathbf{1}_{n_i}\otimes\boldsymbol{x}(k))\|^2\\
&\leq\sum_{j=1}^{n_i}\left(l_{ij}^2\|\boldsymbol{s}_{ij}(k)-\boldsymbol{x}(k)\|^2\right)=\max_{ij\in\mathcal{V}_i}\{l_{ij}^2\}\|\boldsymbol{e}_{s_i}(k)\|^2,
\end{aligned}\tag{C.3}$$

and

$$\begin{aligned}
&\left\|\mathbf{Q}_i(\mathbf{1}_{n_i}\otimes\boldsymbol{x}(k+1))-\mathbf{Q}_i(\mathbf{1}_{n_i}\otimes\boldsymbol{x}(k))\right\|^2\\
&\leq\Big(\sum_{j=1}^{n_i}l_{ij}^2\Big)\left\|\boldsymbol{x}(k+1)-\boldsymbol{x}(k)\right\|^2.
\end{aligned}\tag{C.4}$$

Then, by Young's inequality again, one has

$$\begin{aligned}
&\Big\|(\hat{\mathcal{R}}_i\otimes I_{n_i})\big(\mathbf{1}_{n_i}\otimes(\mathbf{Q}_i(\boldsymbol{s}_i(k)))-\mathbf{Q}_i(\mathbf{1}_{n_i}\otimes\boldsymbol{x}(k))\big)\\
&-\mathbf{1}_{n_i}\otimes\Big(\mathbf{Q}_i(\mathbf{1}_{n_i}\otimes\boldsymbol{x}(k+1))-\mathbf{Q}_i(\mathbf{1}_{n_i}\otimes\boldsymbol{x}(k))\Big)\Big\|^2\\
&\leq2\left(n_i\|\hat{\mathcal{R}}_i\|^2\left\|\mathbf{Q}_i(\boldsymbol{s}_i(k))-\mathbf{Q}_i(\mathbf{1}_{n_i}\otimes\boldsymbol{x}(k))\right\|^2\right.\\
&\quad\left.+n_i\left\|\mathbf{Q}_i(\mathbf{1}_{n_i}\otimes\boldsymbol{x}(k+1))-\mathbf{Q}_i(\mathbf{1}_{n_i}\otimes\boldsymbol{x}(k))\right\|^2\right)\\
&\leq2n_i\left\|\hat{\mathcal{R}}_i\right\|^2\max_{ij\in\mathcal{V}_i}\{l_{ij}^2\}\|\boldsymbol{e}_{\boldsymbol{s}_i}(k)\|^2\\
&\quad+2n_i\sum_{j=1}^{n_i}l_{ij}^2\left\|\beta\hat{\mathcal{L}}^o\check{\mathcal{L}}^o\hat{I}\boldsymbol{\eta}(k)\right\|^2,
\end{aligned}\tag{C.5}$$

where (C.3), (5) and (C.4) are used in the last step. Substituting (C.5) into (C.2) yields

$$\begin{aligned}
&V_\eta(k+1)-V_\eta(k)\\
&\leq-\frac{1}{2}\left\|\boldsymbol{e}_\eta(k)\right\|^2+\max_{i\in\mathcal{I}}\Big\{b_i\|\hat{\mathcal{R}}_i\|^2\max_{ij\in\mathcal{V}_i}\{l_{ij}^2\}\Big\}\|\boldsymbol{e}_s(k)\|^2\\
&+\sum_{i=1}^{N}\Big(b_i\sum_{j=1}^{n_i}l_{ij}^2\Big)\|\alpha\hat{\mathcal{L}}^o\check{\mathcal{L}}^o\hat{I}\boldsymbol{\eta}(k)\|^2,
\end{aligned}\tag{C.6}$$

where $b_i$ is defined in Theorem 1.

Define $V_x(k)=\sum_{i=1}^{N}V_{x_i}(k)$ with $V_{x_i}(k)=\left\|\mathcal{L}_i^{o\mathrm{T}}\frac{\partial f_i}{\partial\boldsymbol{x}_i}(\boldsymbol{x}(k))\right\|^2$. Recalling Lemma 1, it is obvious that $V_x(k)=0$ if and only if $\boldsymbol{x}(k)=\boldsymbol{x}^*$. By definition, one can easily derive that

$$\begin{aligned}
&V_{x_i}(k+1)-V_{x_i}(k)\\
&=\left\|\mathcal{L}_i^{o\mathrm{T}}\frac{\partial f_i}{\partial\boldsymbol{x}_i}(\boldsymbol{x}(k+1))-\mathcal{L}_i^{o\mathrm{T}}\frac{\partial f_i}{\partial\boldsymbol{x}_i}(\boldsymbol{x}(k))\right\|^2\\
&\quad+2\left(\mathcal{L}_i^{o\mathrm{T}}\frac{\partial f_i}{\partial\boldsymbol{x}_i}(\boldsymbol{x}(k+1))-\mathcal{L}_i^{o\mathrm{T}}\frac{\partial f_i}{\partial\boldsymbol{x}_i}(\boldsymbol{x}(k))\right)^{\mathrm{T}}\\
&\quad\times\mathcal{L}_i^{o\mathrm{T}}\frac{\partial f_i}{\partial\boldsymbol{x}_i}(\boldsymbol{x}(k)).
\end{aligned}\tag{C.7}$$

Noting from the definition of $\check{\mathcal{L}}_i^o$ in (4) that

$$\begin{aligned}
&\check{\mathcal{L}}_i^o(I_{n_i}\otimes(\mathbf{1}_{n_i}^T\otimes I_{n_i}))\boldsymbol{\eta}_i(k)-\mathcal{L}_i^{o\mathrm{T}}\frac{\partial f_i}{\partial\boldsymbol{x}_i}(\boldsymbol{x}(k))\\
&=\check{\mathcal{L}}_i^o(I_{n_i}\otimes(\mathbf{1}_{n_i}^T\otimes I_{n_i}))\boldsymbol{\eta}_i(k)-\check{\mathcal{L}}_i^o\Big(\mathbf{1}_{n_i}\otimes\frac{\partial f_i}{\partial\boldsymbol{x}_i}(\boldsymbol{x}(k))\Big)\\
&=\check{\mathcal{L}}_i^o(I_{n_i}\otimes(\mathbf{1}_{n_i}^T\otimes I_{n_i}))\boldsymbol{e}_{\eta_i}(k),
\end{aligned}\tag{C.8}$$

one can derive that

$$\begin{aligned}
&2\left(\mathcal{L}_i^{o\mathrm{T}}\frac{\partial f_i}{\partial\boldsymbol{x}_i}(\boldsymbol{x}(k+1))-\mathcal{L}_i^{o\mathrm{T}}\frac{\partial f_i}{\partial\boldsymbol{x}_i}(\boldsymbol{x}(k))\right)^{\mathrm{T}}\\
&\times\mathcal{L}_i^{o\mathrm{T}}\frac{\partial f_i}{\partial\boldsymbol{x}_i}(\boldsymbol{x}(k))\\
&=2\left(\mathcal{L}_i^{o\mathrm{T}}\frac{\partial f_i}{\partial\boldsymbol{x}_i}(\boldsymbol{x}(k+1))-\mathcal{L}_i^{o\mathrm{T}}\frac{\partial f_i}{\partial\boldsymbol{x}_i}(\boldsymbol{x}(k))\right)^{\mathrm{T}}\\
&\quad\times\Big(\check{\mathcal{L}}_i^o(I_{n_i}\otimes(\mathbf{1}_{n_i}^T\otimes I_{n_i}))\boldsymbol{\eta}_i(k)\\
&\qquad-\check{\mathcal{L}}_i^o(I_{n_i}\otimes(\mathbf{1}_{n_i}^T\otimes I_{n_i}))\boldsymbol{e}_{\eta_i}(k)\Big)\\
&\leq-\frac{2}{\alpha}\left(\frac{\partial f_i}{\partial\boldsymbol{x}_i}(\boldsymbol{x}(k+1))-\frac{\partial f_i}{\partial\boldsymbol{x}_i}(\boldsymbol{x}(k))\right)^{\mathrm{T}}\\
&\quad\times\big(\boldsymbol{x}_i(k+1)-\boldsymbol{x}_i(k)\big)+\left\|\check{\mathcal{L}}_i^o(I_{n_i}\otimes(\mathbf{1}_{n_i}^T\otimes I_{n_i}))\boldsymbol{e}_{\eta_i}(k)\right\|^2\\
&\quad+\left\|\mathcal{L}_i^{o\mathrm{T}}\frac{\partial f_i}{\partial\boldsymbol{x}_i}(\boldsymbol{x}(k+1))-\mathcal{L}_i^{o\mathrm{T}}\frac{\partial f_i}{\partial\boldsymbol{x}_i}(\boldsymbol{x}(k))\right\|^2,
\end{aligned}\tag{C.9}$$

where (C.8) is used in the first step, (5) is used to derived the first term in the last step, and Young's inequality is used to obtain the second and third terms in the last step.

Combining (C.7) and (C.9) yields

$$\begin{aligned}
&V_x(k+1)-V_x(k)\\
&\leq-\frac{2}{\alpha}(\mathbf{P}(\boldsymbol{x}(k+1))-\mathbf{P}(\boldsymbol{x}(k)))^{\mathrm{T}}(\boldsymbol{x}(k+1)-\boldsymbol{x}(k))
\end{aligned}$$

$$
\begin{aligned}
&+\sum_{i=1}^{N}\left\|\breve{\mathcal{L}}_i^{\rm o}(I_{n_i}\otimes(\mathbf{1}_{n_i}^T\otimes I_{n_i}))\boldsymbol{e}_{\eta_i}(k)\right\|^2\\
&+2\sum_{i=1}^{N}\left\|\mathcal{L}_i^{\rm oT}\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k+1))-\mathcal{L}_i^{\rm oT}\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k))\right\|^2\\
\le&-\frac{2\mu}{\alpha}\|\boldsymbol{x}(k+1)-\boldsymbol{x}(k)\|^2+\sum_{i=1}^{N}n_i\left\|\breve{\mathcal{L}}_i^{\rm o}\right\|^2\left\|\boldsymbol{e}_{\eta_i}(k)\right\|^2\\
&+2\sum_{i=1}^{N}\left(l_i^2\left\|\mathcal{L}_i^{\rm o}\right\|^2\|\boldsymbol{x}(k+1)-\boldsymbol{x}(k)\|^2\right)\\
\le&-2\left(\frac{\mu}{\alpha}-\sum_{i=1}^{N}l_i^2\|\mathcal{L}_i^{\rm o}\|^2\right)\left\|\alpha\hat{\mathcal{L}}^{\rm o}\breve{\hat{\mathcal{L}}}^{\rm o}\hat{I}\boldsymbol{\eta}(k)\right\|^2\\
&+\max_{i\in\mathcal{I}}\left\{n_i\left\|\breve{\mathcal{L}}_i^{\rm o}\right\|^2\right\}\left\|\boldsymbol{e}_{\eta}(k)\right\|^2,
\end{aligned}\tag{C.10}
$$

where Assumptions 2 and 3 are used in the second step, and (5) is used in the last step.

Consider the Lyapunov function $V(k)=V_x(k)+\gamma_\eta V_\eta(k)+\gamma_s V_s(k)$, where $\gamma_\eta$ and $\gamma_s$ have been given in Theorem 1 . Combining (C.1), (C.6), (C.10) and the condition (10) yields

$$
\begin{aligned}
&V(k+1)-V(k)\\
&\le-\frac{\gamma_\eta}{4}\left\|\boldsymbol{e}_{\eta}(k)\right\|^2-\frac{\gamma_s}{4}\|\boldsymbol{e}_s(k)\|^2-\mu\alpha\left\|\hat{\mathcal{L}}^{\rm o}\breve{\hat{\mathcal{L}}}^{\rm o}\hat{I}\boldsymbol{\eta}(k)\right\|^2.
\end{aligned}
$$

Note that

$$
\begin{aligned}
&-\left\|\hat{\mathcal{L}}^{\rm o}\breve{\hat{\mathcal{L}}}^{\rm o}\hat{I}\boldsymbol{\eta}(k)\right\|^2=-\sum_{i=1}^{N}\left\|\mathcal{L}_i^{\rm o}\breve{\mathcal{L}}_i^{\rm o}(I_{n_i}\otimes(\mathbf{1}_{n_i}^T\otimes I_{n_i}))\boldsymbol{\eta}_i(k)\right\|^2\\
&=-\sum_{i=1}^{N}\left\|\mathcal{L}_i^{\rm o}\breve{\mathcal{L}}_i^{\rm o}(I_{n_i}\otimes(\mathbf{1}_{n_i}^T\otimes I_{n_i}))\boldsymbol{e}_{\eta_i}(k)+\mathcal{L}_i^{\rm o}\mathcal{L}_i^{\rm oT}\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k))\right\|^2\\
&\le\sum_{i=1}^{N}\left(\left\|\mathcal{L}_i^{\rm o}\breve{\mathcal{L}}_i^{\rm o}(I_{n_i}\otimes(\mathbf{1}_{n_i}^T\otimes I_{n_i}))\boldsymbol{e}_{\eta_i}(k)\right\|^2\right.\\
&\qquad\left.-\frac{1}{2}\left\|\mathcal{L}_i^{\rm o}\mathcal{L}_i^{\rm oT}\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k))\right\|^2\right)\\
&\le\max_{i\in\mathcal{I}}\{n_i\left\|\mathcal{L}_i^{\rm o}\breve{\mathcal{L}}_i^{\rm o}\right\|^2\}\|\boldsymbol{e}_\eta(k)\|^2-\frac{1}{2}\min_{i\in\mathcal{I}}\{\lambda_2(\mathcal{L}_i^{\rm oT}\mathcal{L}_i^{\rm o})\}V_x(k),
\end{aligned}
$$

where (C.8) is used in the second step, the second last inequality is obtained by noticing

$$
\begin{aligned}
&-2\left(\mathcal{L}_i^{\rm o}\breve{\mathcal{L}}_i^{\rm o}(I_{n_i}\otimes(\mathbf{1}_{n_i}^T\otimes I_{n_i}))\boldsymbol{e}_{\eta_i}(k)\right)^T\mathcal{L}_i^{\rm o}\mathcal{L}_i^{\rm oT}\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k))\\
&\le2\left\|\mathcal{L}_i^{\rm o}\breve{\mathcal{L}}_i^{\rm o}(I_{n_i}\otimes(\mathbf{1}_{n_i}^T\otimes I_{n_i}))\boldsymbol{e}_{\eta_i}(k)\right\|^2+\frac{1}{2}\left\|\mathcal{L}_i^{\rm o}\mathcal{L}_i^{\rm oT}\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k))\right\|^2,
\end{aligned}
$$

and the last inequality is derived based on Lemma 2, and $\lambda_2(\mathcal{L}_i^{\rm oT}\mathcal{L}_i^{\rm o})$ denotes the smallest non-zero eigenvalue of the matrix $\mathcal{L}_i^{\rm oT}\mathcal{L}_i^{\rm o}$. Combing the above formulas and condition (14) yields

$$
\begin{aligned}
&V(k+1)-V(k)\\
&\le-\frac{\gamma_\eta}{8}\left\|\boldsymbol{e}_{\eta}(k)\right\|^2-\frac{\gamma_s}{4}\|\boldsymbol{e}_s(k)\|^2\\
&\quad-\frac{\mu\alpha}{2}\min_{i\in\mathcal{I}}\{\lambda_2(\mathcal{L}_i^{\rm oT}\mathcal{L}_i^{\rm o})\}V_x(k)\\
&\le-\epsilon V(k),
\end{aligned}
$$

where $\epsilon=\min\left\{\frac{\mu\alpha}{2}\min_{i\in\mathcal{I}}\{\lambda_2(\mathcal{L}_i^{\rm oT}\mathcal{L}_i^{\rm o})\},\frac{1}{8\|\mathcal{X}_{\mathcal{R}}\|},\frac{1}{4\|\mathcal{X}_{\mathcal{W}}\|}\right\}$. The above inequality shows that $V(k)$ decreases to zero at a linear rate of $\mathcal{O}((1-\epsilon)^k)$. This implies that $\boldsymbol{e}_\eta(k)$, $\boldsymbol{e}_s(k)$, and $\mathrm{col}(\mathcal{L}_1^{\rm oT}\frac{\partial f_1}{\partial \boldsymbol{x}_1}(\boldsymbol{x}(k)),\ldots,\mathcal{L}_N^{\rm o\,T}\frac{\partial f_N}{\partial \boldsymbol{x}_N}(\boldsymbol{x}(k)))$ also linearly converges to zero. Recalling Lemma 1 and the fact that $\boldsymbol{x}(k)\in\Omega,\forall k\in\mathbb{N}$, one can verify that $\boldsymbol{x}(k)$ converges to $\boldsymbol{x}^*$.

To prove the linear convergence of $\boldsymbol{x}(k)$ to $\boldsymbol{x}^*$, one can take note of Eq. (5) and the definition of $\boldsymbol{e}_\eta$. This yields $\boldsymbol{x}(k+1)-\boldsymbol{x}(k)=-\alpha\hat{\mathcal{L}}^{\rm o}\breve{\hat{\mathcal{L}}}^{\rm o}\hat{I}\boldsymbol{e}_\eta(k)-\alpha\hat{\mathcal{L}}^{\rm o}\mathrm{col}(\mathcal{L}_1^{\rm oT}\frac{\partial f_1}{\partial \boldsymbol{x}_1}(\boldsymbol{x}(k)),\ldots,\mathcal{L}_N^{\rm o\,T}\frac{\partial f_N}{\partial \boldsymbol{x}_N}(\boldsymbol{x}(k)))$. Since $\boldsymbol{e}_\eta(k)$ and $\mathrm{col}(\mathcal{L}_1^{\rm oT}\frac{\partial f_1}{\partial \boldsymbol{x}_1}(\boldsymbol{x}(k)),\ldots,\mathcal{L}_N^{\rm o\,T}\frac{\partial f_N}{\partial \boldsymbol{x}_N}(\boldsymbol{x}(k)))$ converge linearly to zero, there exist constants $m>0$ and $0<p<1$ such that $\|\boldsymbol{x}(k+1)-\boldsymbol{x}(k)\|\le mp^k$ for all $k\in\mathbb{N}$. Moreover, since $\boldsymbol{x}(\infty)=\boldsymbol{x}^*$, one has $\|\boldsymbol{x}(k)-\boldsymbol{x}^*\|=\|\boldsymbol{x}(k)-\boldsymbol{x}(k+1)+\boldsymbol{x}(k+1)-\boldsymbol{x}(k+2)+\cdots-\boldsymbol{x}(\infty)\|\le mp^k\sum_{h=0}^{\infty}p^h=\frac{m}{1-p}p^k$. Then, one can conclude that the collective state $\boldsymbol{x}$ linearly converges to the NE $\boldsymbol{x}^*$.

## Appendix D. Proof of Theorem 2

Similar to (5), one can obtain from (16a) and (16b) that

$$
\begin{aligned}
&\boldsymbol{x}_i(k+1)-\boldsymbol{x}_i(k)=-\beta\mathcal{L}_i^{\rm o}\breve{\mathcal{L}}_i^{\rm o}\boldsymbol{\xi}_i(k),\\
&\boldsymbol{x}(k+1)-\boldsymbol{x}(k)=-\beta\hat{\mathcal{L}}^{\rm o}\breve{\hat{\mathcal{L}}}^{\rm o}\boldsymbol{\xi}(k),
\end{aligned}\tag{D.1}
$$

where $\hat{\mathcal{L}}^{\rm o}$ and $\breve{\hat{\mathcal{L}}}^{\rm o}$ are defined below (5).

Consider the function $V_s(k)$ defined before (C.1). Similar to (C.1), one can derive from (15) that

$$
\begin{aligned}
&V_s(k+1)-V_s(k)\\
&\le-\frac{1}{2}\|\boldsymbol{e}_s(k)\|^2+n_{\rm sum}b\left\|\beta\hat{\mathcal{L}}^{\rm o}\breve{\hat{\mathcal{L}}}^{\rm o}\boldsymbol{\xi}(k)\right\|^2.
\end{aligned}\tag{D.2}
$$

Define $V_\xi(k)=\boldsymbol{e}_\xi(k)^T\mathcal{X}_{\mathcal{C}}\boldsymbol{e}_\xi(k)$, where $\mathcal{X}_{\mathcal{C}}=\mathrm{diag}\{\mathcal{X}_{\mathcal{C}_1}\otimes I_{n_1},\ldots,\mathcal{X}_{\mathcal{C}_N}\otimes I_{n_N}\}$. Similar to (C.2), one can obtain from (20) that

$$
\begin{aligned}
&V_\xi(k+1)-V_\xi(k)\\
&\le-\frac{1}{2}\|\boldsymbol{e}_\xi(k)\|^2+\sum_{i=1}^{N}\left(\left(2\|\bar{\mathcal{C}}_i^T\mathcal{X}_{\mathcal{C}i}\bar{I}_i\|^2+\|\bar{I}_i^T\mathcal{X}_{\mathcal{C}i}\bar{I}_i\|\right)\right.\\
&\qquad\left.\times\|\mathbf{Q}_i(\boldsymbol{s}_i(k+1))-\mathbf{Q}_i(\boldsymbol{s}_i(k))\|^2\right).
\end{aligned}\tag{D.3}
$$

Under Assumption 2, one has

$$
\begin{aligned}
&\|\mathbf{Q}_i(\boldsymbol{s}_i(k+1))-\mathbf{Q}_i(\boldsymbol{s}_i(k))\|^2\\
&=\sum_{j=1}^{n_i}\left\|\frac{\partial f_{ij}}{\partial \boldsymbol{x}_i}(\boldsymbol{s}_{ij}(k+1))-\frac{\partial f_{ij}}{\partial \boldsymbol{x}_i}(\boldsymbol{s}_{ij}(k))\right\|^2\\
&\le\max_{ij\in\mathcal{V}_i}\{l_{ij}^2\}\left\|\boldsymbol{s}_i(k+1)-\boldsymbol{s}_i(k)\right\|^2.
\end{aligned}\tag{D.4}
$$

From (16d), one has

$$
\begin{aligned}
&\boldsymbol{s}(k+1)-\boldsymbol{s}(k)\\
&=(\mathcal{W}\otimes I_{n_{\rm sum}}+\bar{\mathcal{W}}-I_{n_{\rm sum}^2})\boldsymbol{s}(k)+\hat{\mathcal{W}}\left(\mathbf{1}_{n_{\rm sum}}\otimes\boldsymbol{x}(k)\right)\\
&=(\mathcal{H}_{\mathcal{W}}-I_{n_{\rm sum}^2})\boldsymbol{e}_s(k),
\end{aligned}\tag{D.5}
$$

where $\mathcal{H}_{\mathcal{W}}=\mathcal{W}\otimes I_{n_{\rm sum}}+\bar{\mathcal{W}}$, and the last equality is obtained by noticing $(\mathcal{W}\otimes I_{n_{\rm sum}}+\bar{\mathcal{W}}+\hat{\mathcal{W}})(\mathbf{1}_{n_{\rm sum}}\otimes\boldsymbol{x})=\mathbf{1}_{n_{\rm sum}}\otimes\boldsymbol{x}$. Substituting (D.4) and (D.5) back into (D.3) yields

$$
\begin{aligned}
&V_\xi(k+1)-V_\xi(k)\\
&\le-\frac{1}{2}\|\boldsymbol{e}_\xi(k)\|^2+\tilde{b}\|I_{n_{\rm sum}^2}-\mathcal{H}_{\mathcal{W}}\|^2\|\boldsymbol{e}_s(k)\|^2,
\end{aligned}\tag{D.6}
$$

where $\tilde{b}$ is given in Theorem 2.

Define $\tilde{V}_x(k)=\sum_{i=1}^{N}\frac{1}{2n_i}\tilde{V}_{x_i}(k)$, where

$$
\tilde{V}_{x_i}(k)=\left(\mathcal{L}_i^{\rm oT}\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k))\right)^T\hat{\boldsymbol{v}}_i\left(\mathcal{L}_i^{\rm oT}\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k))\right).
$$

One can derive that

$$
\tilde{V}_{x_i}(k+1)-\tilde{V}_{x_i}(k)
$$

$$
\begin{aligned}
&= \left(\mathcal{L}_i^{oT}\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k+1)) - \mathcal{L}_i^{oT}\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k))\right)^T \hat{\boldsymbol{v}}_i \\
&\quad \times \left(\mathcal{L}_i^{oT}\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k+1)) - \mathcal{L}_i^{oT}\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k))\right) \\
&\quad + 2\Big(\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k+1)) - \frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k))\Big)^T \mathcal{L}_i^o \hat{\boldsymbol{v}}_i \mathcal{L}_i^{oT}\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k)).
\end{aligned} \tag{D.7}
$$

Note from the definition of $\check{\mathcal{L}}_i^o$ that

$$
\check{\mathcal{L}}_i^o \boldsymbol{e}_{\xi_i} = \check{\mathcal{L}}_i^o \boldsymbol{\xi}_i - \check{\mathcal{L}}_i^o(\boldsymbol{v}_i \otimes \bar{\boldsymbol{\xi}}_i) = \check{\mathcal{L}}_i^o \boldsymbol{\xi}_i - \hat{\boldsymbol{v}}_i \mathcal{L}_i^{oT}\bar{\boldsymbol{\xi}}_i. \tag{D.8}
$$

Combing (18) and (D.8), one has

$$
\begin{aligned}
\hat{\boldsymbol{v}}_i \mathcal{L}_i^{oT}\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k)) &= n_i \check{\mathcal{L}}_i^o \boldsymbol{\xi}_i(k) - n_i \check{\mathcal{L}}_i^o \boldsymbol{e}_{\xi_i}(k) \\
&\quad + \hat{\boldsymbol{v}}_i \mathcal{L}_i^{oT}\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k)) - n_i \hat{\boldsymbol{v}}_i \mathcal{L}_i^{oT}\bar{\mathbf{Q}}_i(\boldsymbol{s}_i(k)).
\end{aligned} \tag{D.9}
$$

It follows from (D.1) and (D.9) that

$$
\begin{aligned}
&2\Big(\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k+1)) - \frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k))\Big)^T \mathcal{L}_i^o \hat{\boldsymbol{v}}_i \mathcal{L}_i^{oT}\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k)) \\
&= -\frac{2n_i}{\beta}\Big(\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k+1)) - \frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k))\Big)^T (\boldsymbol{x}_i(k+1) - \boldsymbol{x}_i(k)) \\
&\quad + 2\Big(\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k+1)) - \frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k))\Big)^T \mathcal{L}_i^o \hat{\boldsymbol{v}}_i \\
&\quad \times \left(\mathcal{L}_i^{oT}\Big(\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k)) - n_i \bar{\mathbf{Q}}_i(\boldsymbol{s}_i(k))\Big) - n_i \hat{\boldsymbol{v}}_i^{-1}\check{\mathcal{L}}_i^o \boldsymbol{e}_{\xi_i}(k)\right).
\end{aligned}
$$

Substituting the above equation into (D.7) and using Young's inequality yield

$$
\begin{aligned}
&\tilde{V}_{x_i}(k+1) - \tilde{V}_{x_i}(k) \\
&\le -\frac{2n_i}{\beta}\Big(\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k+1)) - \frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k))\Big)^T (\boldsymbol{x}_i(k+1) - \boldsymbol{x}_i(k)) \\
&\quad + 2\|\hat{\boldsymbol{v}}_i\|\|\mathcal{L}_i^o\|^2 \left\|\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k+1)) - \frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k))\right\|^2 \\
&\quad + 2\|\hat{\boldsymbol{v}}_i\|\|\mathcal{L}_i^o\|^2 \left\|\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k)) - n_i \bar{\mathbf{Q}}_i(\boldsymbol{s}_i(k))\right\|^2 \\
&\quad + 2\|\hat{\boldsymbol{v}}_i^{-1}\| \left\|n_i \check{\mathcal{L}}_i^o \boldsymbol{e}_{\xi_i}(k)\right\|^2.
\end{aligned} \tag{D.10}
$$

Under Assumption 2,

$$
\begin{aligned}
&\left\|n_i \bar{\mathbf{Q}}_i(\boldsymbol{s}_i(k)) - \frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k))\right\|^2 \\
&= \left\|\sum_{j=1}^{n_i}\left(\frac{\partial f_{ij}}{\partial \boldsymbol{x}_i}(\boldsymbol{s}_{ij}(k)) - \frac{\partial f_{ij}}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k))\right)\right\|^2 \\
&\le \left(\sum_{j=1}^{n_i} l_{ij}\|\boldsymbol{s}_{ij}(k) - \boldsymbol{x}(k)\|\right)^2 \\
&\le \Big(\sum_{j=1}^{n_i} l_{ij}^2\Big)\Big(\sum_{j=1}^{n_i}\|\boldsymbol{s}_{ij}(k) - \boldsymbol{x}(k)\|^2\Big).
\end{aligned} \tag{D.11}
$$

Combining (D.10), (D.11), (D.1) and Assumptions 2 and 3 yields

$$
\begin{aligned}
&\tilde{V}_x(k+1) - \tilde{V}_x(k) \\
&\le -\Big(\frac{\mu}{\beta} - \sum_{i=1}^{N}\frac{\|\hat{\boldsymbol{v}}_i\|\,\|\mathcal{L}_i^o\|^2 l_i^2}{n_i}\Big)\left\|\beta \hat{\mathcal{L}}^o \check{\mathcal{L}}^o \boldsymbol{\xi}(k)\right\|^2 \\
&\quad + \max_{i\in\mathcal{I}}\left\{\frac{\|\hat{\boldsymbol{v}}_i\|\,\|\mathcal{L}_i^o\|^2 \sum_{j=1}^{n_i} l_{ij}^2}{n_i}\right\}\|\boldsymbol{e}_s(k)\|^2 \\
&\quad + \max_{i\in\mathcal{I}}\{n_i\|\hat{\boldsymbol{v}}_i^{-1}\|\|\check{\mathcal{L}}_i^o\|^2\}\left\|\boldsymbol{e}_\xi(k)\right\|^2.
\end{aligned} \tag{D.12}
$$

Consider the Lyapunov function $\tilde{V}(k) = \tilde{V}_x(k) + \gamma_\xi V_\xi(k) + \tilde{\gamma}_s V_s(k)$, where $\gamma_\xi$ and $\tilde{\gamma}_s$ are defined in Theorem 2. By (D.2), (D.6), (D.12) and condition (21), one has

$$
\begin{aligned}
&\tilde{V}(k+1) - \tilde{V}(k) \\
&\le -\frac{\mu\beta}{2}\left\|\hat{\mathcal{L}}^o \check{\mathcal{L}}^o \boldsymbol{\xi}(k)\right\|^2 - \frac{\gamma_\xi}{4}\|\boldsymbol{e}_\xi(k)\|^2 - \frac{\tilde{\gamma}_s}{4}\|\boldsymbol{e}_s(k)\|^2.
\end{aligned} \tag{D.13}
$$

Note that

$$
\begin{aligned}
&-\left\|\hat{\mathcal{L}}^o \check{\mathcal{L}}^o \boldsymbol{\xi}(k)\right\|^2 = -\sum_{i=1}^{N}\left\|\mathcal{L}_i^o \check{\mathcal{L}}_i^o \boldsymbol{\xi}_i(k)\right\|^2 \\
&= -\sum_{i=1}^{N}\left\|\mathcal{L}_i^o\left(\check{\mathcal{L}}_i^o \boldsymbol{e}_{\xi_i}(k) + \hat{\boldsymbol{v}}_i \mathcal{L}_i^{oT}\Big(\bar{\mathbf{Q}}_i(\boldsymbol{s}_i(k))\right.\right. \\
&\quad \left.\left.- \frac{1}{n_i}\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k))\Big)\right) + \frac{1}{n_i}\mathcal{L}_i^o \hat{\boldsymbol{v}}_i \mathcal{L}_i^{oT}\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k))\right\|^2 \\
&\le \sum_{i=1}^{N}\left(2\left\|\mathcal{L}_i^o \check{\mathcal{L}}_i^o\right\|^2\left\|\boldsymbol{e}_{\xi_i}(k)\right\|^2\right. \\
&\quad + 2\frac{\|\mathcal{L}_i^o \hat{\boldsymbol{v}}_i \mathcal{L}_i^{oT}\|^2}{n_i^2}\left\|n_i \bar{\mathbf{Q}}_i(\boldsymbol{s}_i(k)) - \frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k))\right\|^2 \\
&\quad \left.- \frac{1}{2n_i^2}\left\|\mathcal{L}_i^o \hat{\boldsymbol{v}}_i \mathcal{L}_i^{oT}\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k))\right\|^2\right),
\end{aligned} \tag{D.14}
$$

where the second equality is obtained by (18) and (D.8), and the last inequality is derived based on Young's inequality.

Under Assumption 1, zero is a simple eigenvalue of $\hat{\boldsymbol{v}}_i^{\frac{1}{2}}\mathcal{L}_i^{oT}\mathcal{L}_i^o \hat{\boldsymbol{v}}_i^{\frac{1}{2}}$, and all other eigenvalues are positive. Denote $\boldsymbol{r}_i$ the right eigenvector of $\mathcal{L}_i^o$ corresponding to the zero eigenvalue, i.e., $\mathcal{L}_i^o \boldsymbol{r}_i = \mathbf{0}$. Obviously, $\hat{\boldsymbol{v}}_i^{-\frac{1}{2}}\boldsymbol{r}_i$ is a right eigenvector of the matrix $(\hat{\boldsymbol{v}}_i^{\frac{1}{2}}\mathcal{L}_i^{oT}\mathcal{L}_i^o \hat{\boldsymbol{v}}_i^{\frac{1}{2}})$ corresponding to its zero eigenvalue. Since $\hat{\boldsymbol{v}}_i^{\frac{1}{2}}\mathcal{L}_i^{oT}\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k)) \perp \hat{\boldsymbol{v}}_i^{-\frac{1}{2}}\boldsymbol{r}_i$, one can derive from Lemma 2 that

$$
\begin{aligned}
&\left\|\mathcal{L}_i^o \hat{\boldsymbol{v}}_i \mathcal{L}_i^{oT}\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k))\right\|^2 = \left(\hat{\boldsymbol{v}}_i^{\frac{1}{2}}\mathcal{L}_i^{oT}\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k))\right)^T \\
&\quad \times \left(\hat{\boldsymbol{v}}_i^{\frac{1}{2}}\mathcal{L}_i^{oT}\mathcal{L}_i^o \hat{\boldsymbol{v}}_i^{\frac{1}{2}}\right)\left(\hat{\boldsymbol{v}}_i^{\frac{1}{2}}\mathcal{L}_i^{oT}\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k))\right) \\
&\ge \lambda_2\left(\hat{\boldsymbol{v}}_i^{\frac{1}{2}}\mathcal{L}_i^{oT}\mathcal{L}_i^o \hat{\boldsymbol{v}}_i^{\frac{1}{2}}\right)\left\|\hat{\boldsymbol{v}}_i^{\frac{1}{2}}\mathcal{L}_i^{oT}\frac{\partial f_i}{\partial \boldsymbol{x}_i}(\boldsymbol{x}(k))\right\|^2.
\end{aligned} \tag{D.15}
$$

Substituting (D.11) and (D.15) into (D.14) yields

$$
\begin{aligned}
&-\left\|\hat{\mathcal{L}}^o \check{\mathcal{L}}^o \boldsymbol{\xi}(k)\right\|^2 \le 2\max_{i\in\mathcal{I}}\{\|\mathcal{L}_i^o \check{\mathcal{L}}_i^o\|^2\}\|\boldsymbol{e}_\xi(k)\|^2 \\
&+ 2\max_{i\in\mathcal{I}}\left\{\frac{\|\mathcal{L}_i^o \hat{\boldsymbol{v}}_i \mathcal{L}_i^{oT}\|^2 \sum_{j=1}^{n_i} l_{ij}^2}{n_i^2}\right\}\|\boldsymbol{e}_s(k)\|^2 \\
&- \min_{i\in\mathcal{I}}\left\{\frac{\lambda_2\left(\hat{\boldsymbol{v}}_i^{\frac{1}{2}}\mathcal{L}_i^{oT}\mathcal{L}_i^o \hat{\boldsymbol{v}}_i^{\frac{1}{2}}\right)}{n_i}\right\}\tilde{V}_x(k).
\end{aligned}
$$

Substituting the above inequality into (D.13) and then using (21) yields

$$
\begin{aligned}
&\tilde{V}(k+1) - \tilde{V}(k) \\
&\le -\frac{\mu\beta}{2}\min_{i\in\mathcal{I}}\left\{\frac{\lambda_2\left(\hat{\boldsymbol{v}}_i^{\frac{1}{2}}\mathcal{L}_i^{oT}\mathcal{L}_i^o \hat{\boldsymbol{v}}_i^{\frac{1}{2}}\right)}{n_i}\right\}\tilde{V}_x(k) \\
&\quad - \Big(\frac{\gamma_\xi}{4} - \mu\beta\max_{i\in\mathcal{I}}\{\|\mathcal{L}_i^o \check{\mathcal{L}}_i^o\|^2\}\Big)\|\boldsymbol{e}_\xi(k)\|^2
\end{aligned}
$$

$$
\begin{aligned}
&-\left(\frac{\tilde{\gamma}_s}{4}-\mu\beta\max_{i\in\mathcal{I}}\left\{\frac{\|\mathcal{L}_i^{o}\hat{\boldsymbol{v}}_i\mathcal{L}_i^{oT}\|^2\sum_{j=1}^{n_i} l_{ij}^2}{n_i^2}\right\}\right)\|\boldsymbol{e}_s(k)\|^2\\
&\leq -\frac{\mu\beta}{2}\min_{i\in\mathcal{I}}\left\{\frac{\lambda_2\left(\hat{\boldsymbol{v}}_i^{\frac{1}{2}}\mathcal{L}_i^{oT}\mathcal{L}_i^{o}\hat{\boldsymbol{v}}_i^{\frac{1}{2}}\right)}{n_i}\right\}\tilde{V}_x(k)\\
&\quad-\frac{\gamma_\xi}{8}\|\boldsymbol{e}_\xi(k)\|^2-\frac{\tilde{\gamma}_s}{8}\|\boldsymbol{e}_s(k)\|^2\\
&\leq -\tilde{\epsilon}\tilde{V}(k),
\end{aligned}
$$

where $\tilde{\epsilon} = \min\left\{\frac{\mu\beta}{2}\min_{i\in\mathcal{I}}\left\{\frac{\lambda_2\left(\hat{\boldsymbol{v}}_i^{\frac{1}{2}}\mathcal{L}_i^{oT}\mathcal{L}_i^{o}\hat{\boldsymbol{v}}_i^{\frac{1}{2}}\right)}{n_i}\right\}, \frac{1}{8\|\mathcal{X}_{\mathcal{C}}\|}, \frac{1}{8\|\mathcal{X}_{\mathcal{W}}\|}\right\}$.

This implies that $\tilde{V}(k)$ converges to the origin at a linear rate of $\mathcal{O}((1-\bar{\epsilon})^k)$. Then, similar to the analysis at the end of the proof of Theorem 1, one can conclude that the collective state $\boldsymbol{x}$ converges to the NE $\boldsymbol{x}^*$ linearly.

## References


Anderson, T., & Martínez, S. (2021). Distributed resource allocation with binary decisions via Newton-like neural network dynamics. *Automatica*, *128*, Article 109564.

Bazaraa, M. S., Sherali, H. D., & Shetty, C. M. (2006). *Nonlinear programming: theory and algorithms* (3rd ed.). John Wiley and Sons, Inc..

Chen, C.-T. (1999). *Linear system theory and design*. Oxford University Press.

Chen, G., & Li, Z. (2018). A fixed-time convergent algorithm for distributed convex optimization in multi-agent systems. *Automatica*, *95*, 539–543.

Deng, Z., & Liu, Y. (2023). Nash equilibrium seeking algorithm design for distributed nonsmooth multicluster games over weight-balanced digraphs. *IEEE Transactions on Neural Networks and Learning Systems*, *34*(12), 10802–10811.

Horn, R. A., Rhee, N. H., & Wasin, S. (1998). Eigenvalue inequalities and equalities. *Linear Algebra and its Applications*, *270*(1), 29–44.

Huang, S., Lei, J., & Hong, Y. (2023). A linearly convergent distributed Nash equilibrium seeking algorithm for aggregative games. *IEEE Transactions on Automatic Control*, *68*(3), 1753–1759.

Lou, Y., Hong, Y., Xie, L., Shi, G., & Johansson, K. H. (2016). Nash equilibrium computation in subnetwork zero-sum games with switching communications. *IEEE Transactions on Automatic Control*, *61*(10), 2920–2935.

Mei, J., Ren, W., & Chen, J. (2016). Distributed consensus of second-order multi-agent systems with heterogeneous unknown inertias and control gains under a directed graph. *IEEE Transactions on Automatic Control*, *61*(8), 2019–2034.

Meng, M., & Li, X. (2023). On the linear convergence of distributed Nash equilibrium seeking for multi-cluster games under partial-decision information. *Automatica*, *151*, Article 110919.

Nedić, A., Olshevsky, A., & Shi, W. (2017). Achieving geometric convergence for distributed optimization over time-varying graphs. *SIAM Journal on Optimization*, *27*(4), 2597–2633.

Nian, X., Niu, F., & Yang, Z. (2022). Distributed Nash equilibrium seeking for multicluster game under switching communication topologies. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *52*(7), 4105–4116.

Pang, Y., & Hu, G. (2022). Nash equilibrium seeking in N-coalition games via a gradient-free method. *Automatica*, *136*, Article 110013.

Pavel, L. (2020). Distributed GNE seeking under partial-decision information over networks via a doubly-augmented operator splitting approach. *IEEE Transactions on Automatic Control*, *65*(4), 1584–1597.

Ren, W., & Beard, R. W. (2005). Consensus seeking in multiagent systems under dynamically changing interaction topologies. *IEEE Transactions on Automatic Control*, *50*(5), 655–661.

Simpson-Porco, J. W., Poolla, B. K., Monshizadeh, N., & Dörfler, F. (2020). Input–output performance of linear-quadratic saddle-point algorithms with application to distributed resource allocation problems. *IEEE Transactions on Automatic Control*, *65*(5), 2032–2045.

Wei, Y., Shang, C., Fang, H., Zeng, X., Dou, L., & Pardalos, P. (2022). Solving a class of nonsmooth resource allocation problems with directed graphs through distributed Lipschitz continuous multi-proximal algorithms. *Automatica*, *136*, Article 110071.

Xiao, L., & Boyd, S. (2006). Optimal scaling of a gradient method for distributed resource allocation. *Journal of Optimization Theory and Applications*, *129*(3), 469–488.

Xin, R., & Khan, U. A. (2020). Distributed heavy-ball: A generalization and acceleration of first-order methods with gradient tracking. *IEEE Transactions on Automatic Control*, *65*(6), 2627–2633.

Yang, T., Yi, X., Wu, J., Yuan, Y., Wu, D., Meng, Z., Hong, Y., Wang, H., Lin, Z., & Johansson, K. H. (2019). A survey of distributed optimization. *Annual Reviews in Control*, *47*, 278–305.

Yi, P., Hong, Y., & Liu, F. (2016). Initialization-free distributed algorithms for optimal resource allocation with feasibility constraints and application to economic dispatch of power systems. *Automatica*, *74*, 259–269.

Yi, P., Lei, J., & Hong, Y. (2018). Distributed resource allocation over random networks based on stochastic approximation. *Systems & Control Letters*, *114*, 44–51.

Zeng, X., Chen, J., Liang, S., & Hong, Y. (2019). Generalized nash equilibrium seeking strategy for distributed nonsmooth multi-cluster game. *Automatica*, *103*, 20–26.

Zhou, J., Lv, Y., Wen, G., Lü, J., & Zheng, D. (2023). Distributed Nash equilibrium seeking in consistency-constrained multicoalition games. *IEEE Transactions on Cybernetics*, *53*(6), 3675–3687.

Zhou, J., Lv, Y., Wen, C., & Wen, G. (2022). Solving specified-time distributed optimization problem via sampled-data-based algorithm. *IEEE Transactions on Network Science and Engineering*, *9*(4), 2747–2758.

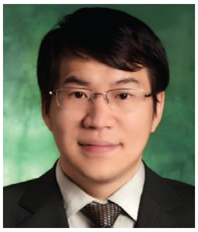
Zhou, J., Wen, G., Lv, Y., Yang, T., & Chen, G. (2024). Distributed resource allocation over multiple interacting coalitions: A game-theoretic approach. *IEEE Transactions on Automatic Control*, *69*(11), 8128–8135.


**Jialing Zhou** received the Ph.D. degree in mechanical systems and control from Peking University, Beijing, China, in 2017. She is currently a Professor with the School of Interdisciplinary Science, Beijing Institute of Technology, Beijing, China. She was selected as a Young Top-Notch Talent of the National Ten Thousand Talents Program and was a recipient of the Young Elite Scientist Sponsorship Program from the China Association for Science and Technology. She serves as an Associate Editor for Asian Journal of Control. Her research interests include distributed control and optimization, guidance and control of flight vehicles, reinforcement learning, and networked games.

**Guanghui Wen** received the Ph.D. degree in mechanical systems and control from Peking University, Beijing, China, in 2012. He is currently an Endowed Chair Professor and the Vice Dean of the School of Automation, Southeast University, Nanjing, China. His current research interests include coordination control of autonomous intelligent systems, distributed resilient control, and reinforcement learning. Prof. Wen was the recipient of the National Science Fund for Distinguished Young Scholars and the Australian Research Council Discovery Early Career Researcher Award. He currently serves as an Associate Editor of the IEEE Transactions on Industrial Informatics, the IEEE Transactions on Control of Network Systems, etc. He is an IET Fellow.

**Yuezu Lv** received the B.S. degree in engineering mechanism and Ph.D. degree in mechanical systems and control from the College of Engineering, Peking University, Beijing, China, in 2013 and 2018, respectively. He is currently a Professor with School of Artificial Intelligence, Beijing Institute of Technology, Beijing, China. His research interests include cooperative control of multi-agent systems, adaptive control, robust control of uncertain systems, and distributed resilient control. He received the 2021 APNNS Young Researcher Award by Asia Pacific Neural Network Society. He is an AE of IEEE SMC Magazine, and a young editorial board member of Artificial Intelligence Science and Engineering.

**Xinlei Yi** received the B.S. and M.S. degrees in mathematics from China University of Geoscience, Wuhan, China, and Fudan University, Shanghai, China, in 2011 and 2014, respectively, and the Ph.D. degree in electrical engineering from KTH Royal Institute of Technology, Stockholm, Sweden, in 2020. He was a Postdoc with KTH Royal Institute of Technology from 2020 to 2022 and with the Lab for Information & Decision Systems, Massachusetts Institute of Technology, Cambridge, MA, USA, from 2022 to 2024. He is a tenure-track professor of Shanghai Institute of Intelligent Science and Technology, Tongji University. Dr. Yi's was selected as one of the four finalists for the 2021 European Systems & Control Ph.D. Thesis Award. His current research interests include distributed online and optimization, online optimization, meta-learning and graph neural networks.



**Tao Yang** is a Professor at the State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University. He was an Assistant Professor at the Department of Electrical Engineering, University of North Texas, Denton, USA, from 2016-2019. He received the Ph.D. degree in electrical engineering from Washington State University in 2012. Between August 2012 and August 2014, he was an ACCESS postdoctoral researcher with the ACCESS Linnaeus Centre, Royal Institute of Technology, Sweden. He then joined the Pacific Northwest National Laboratory as a postdoc, and was promoted to Scientist/Engineer II in 2015. His research interests include industrial artificial intelligence, integrated optimization and control, distributed control and optimization with applications to process industries, cyber physical systems, networked control systems, and multi-agent systems. He is an Associate Editor for IEEE Transactions on Control of Network Systems, IEEE Transactions on Control Systems Technology, and IEEE Transactions on Neural Networks and Learning System. He received Ralph E. Powe Junior Faculty Enhancement Award and Best Student Paper award (as an advisor) of several international Conference.

**Karl H. Johansson** is Swedish Research Council Distinguished Professor in Electrical Engineering and Computer Science at KTH Royal Institute of Technology in Sweden and Director of Digital Futures. He received M.Sc. degree in Electrical Engineering and Ph.D. in Automatic Control from Lund University. He has held visiting positions at UC Berkeley, Caltech, NTU among other institutions. His research interests are in networked control systems and cyber–physical systems with applications in transportation, energy, and automation networks. He is Vice President IEEE Control Systems Society, member of IFAC Council, and Past President of the European Control Association, and has served on Swedish Scientific Council for Natural Sciences and Engineering Sciences. He has received several best paper awards and other distinctions from IEEE, IFAC, and ACM. He has been awarded Distinguished Professor by Swedish Research Council, Wallenberg Scholar with the Knut and Alice Wallenberg Foundation, Future Research Leader from the Swedish Foundation for Strategic Research, the triennial IFAC Young Author Prize, IEEE Control Systems Society Distinguished Lecturer, and IEEE Control Systems Society Hendrik W. Bode Lecture Prize. He is Fellow of the IEEE and the Royal Swedish Academy of Engineering Sciences.