Nash equilibrium of multi-agent graphical game with a privacy information encrypted learning algorithm

Kun Zhang $^{a,b}$, Ji-Feng Zhang $^{a,c}$, Rong Su $^b$, Huaguang Zhang $^d$

$^a$ Key Laboratory of Systems and Control, Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Science, Beijing 100190, China

$^b$ School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798, Singapore

$^c$ School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

$^d$ School of Information Science and Engineering, Northeastern University, Shenyang 110819, China

Abstract

This paper studies the global Nash equilibrium problem of leader-follower multi-agent dynamics, which yields consensus with a privacy information encrypted learning algorithm. With the secure hierarchical structure, the relationship between the secure consensus problem and global Nash equilibrium is discussed under potential packet loss attacks, and the necessary and sufficient condition for the existence of global Nash equilibrium is provided regarding the soft-constrained graphical game. To achieve the optimal policies, the convergence of decentralized learning algorithm is guaranteed with an iteratively updated pair of decoupled gains. By using the developed quantization scheme and additive-multiplicative property, the encryption-decryption is successfully embedded in the data transmission and computation to overcome the potential privacy violation in unreliable networks. A simulation example is provided to verify the effectiveness of the designed algorithm.

Key words: Networked control systems; graphical game; Nash equilibrium; iteration learning; network security

1 Introduction

Game theory depicts a lot of social scenes in, e.g., economic, militarily, engineering fields, where the decision makers in a multiple-player game will attempt to take actions based on the available information to maximize their utility levels. Because no absolutely complete information of opponent or cooperator can be accessed, there are always uncertainties about the strategies of other partners, no matter how they struggle to off-set the resistances of uncertainties by enhancing their policies. Nash equilibrium can provide a strategic advantage for players, which serves as the core basis of behavior judgment for policy making and action direction up to now. Compared with the hard-bounded game (Van Den Broek et al., 2003), the soft-constrained game, such as zero-sum game, can provide an ideal environment to study multi-player decision and control problems, and has been successfully developed strategic behavior for some complex dynamics (Zhang and Guo, 2019; Zhang et al., 2011; Abu-Khalaf et al., 2006; Lopez et al., 2020). However, the global Nash equilibrium of a game is very difficult to be achieved in the real world, even in a theory, due to the multi-dimensional multi-player networked utility relationships.

Multi-agent systems have been extensively studied over the past two decades, which cover lots of the fundamental cooperative control problems (Ma and Zhang, 2010; Gu et al., 2011; Liu and Huang, 2021; Gao et al., 2018; Liu and Jiang, 2014). For the consensus problem of continuous-time linear multi-agent systems, it is required that the dynamics should be stabilizable and...
detectable, and the undirected communication topology is connected in (Ma and Zhang, 2010). With the designed dynamic filter (Gu et al., 2011) of discrete-time agent systems, the consensus ability condition of single-input agent dynamics is relaxed. Robust synchronization of multi-agent networks is considered in (Trentelman et al., 2013), where the perturbations are assumed to be stable and bounded in $H_{\infty}$-norm by some a priori given tolerance. The effects of undirected graphs on consensus ability and optimal convergence rate are quantified in (You and Xie, 2011), where an extension of directed graph can also be found. A general class of distributed control laws is proposed by (Jiang and Jiang, 2020) for leader-following consensus over jointly connected switching digraphs under some conditions. These work successfully settled the foundation and condition to achieve the consensus, but the global Nash equilibrium in multi-agent systems has few research results.

Information security is also an important topic and concern in the networked control systems. Since the human-cyber-physical system is originated in the manufacturing production (Zhu and Basar, 2015), many privacy information has been exposed to the spiteful network users unintentionally, and packet loss from unreliable communication channels has frequented. Along with the development, some sensitive components or confidential data, such as economic policy, military activity or business relations, will cause enormous property damage and economic loss, if they are tampered with or deleted by the potential attackers. For the data losses problem, the control and estimation problems are studied in (Lin et al., 2017; Zhang et al., 2015; Imer et al., 2006; Lin et al., 2015; Sinopoli et al., 2004) by modeling the arrival of observation in unreliable communication channels as a random process. A necessary and sufficient consensus condition is discussed in (Xu et al., 2019) for multi agents with independent and identically distributed channel losses. However, compared with these recognizable data changes, the unknown leakage of confidential and private information plays a more damaging role in the systems.

To deal with the situation of unknown information leakage, privacy preserving approaches have recently been proposed for the cloud computing, such as differential privacy methods (Wang et al., 2017), adding artificial noise (Mo and Murray, 2016), etc. A distributed optimization algorithm that preserves differential privacy is developed in (Han et al., 2016), which can guarantee user privacy regardless of any auxiliary information an adversary may have. $\epsilon$-differential privacy is introduced in (Liang et al., 2020) to solve the Knapsack problem, where a greedy strategy is proposed. Especially, homomorphic encryption techniques provide an efficient data operation or computation method with partial arithmetic rules (Hadjicostis and Domínguez-García, 2020; Lu and Zhu, 2018; Kim et al., 2016; Kogiso and Fujita, 2015), that are either additive, or multiplicative, or even fully homomorphic. A cloud based model predictive control scheme is discussed with semi-honest servers in (Alexandru et al., 2018), where the two-party privacy is proposed, and that consists of a client-server model and a two-server model. Paillier’s encryption method is developed by (Murguia et al., 2020) into a control scheme for linear time-invariant systems, and achieves a semi-homomorphic encrypted control algorithm. These privacy preserving approaches construct a new control architecture for networked dynamics, that, however, brings lots of secure technical problems for differential systems.

In this paper, we study the consensus of multi-agent graphical dynamics under a set of soft-constrained game, and solve the global Nash equilibrium with a privacy information encrypted learning algorithm. The main contributions can be summarized as follows.

- We have analyzed the consensus problem of leader-follower systems with a soft-constrained game in this paper, where the necessary and sufficient condition for the existence of global Nash equilibrium is obtained. To our best knowledge, no global Nash equilibrium has been considered from such an overall perspective of multi-agent systems.
- This work attempts to solve the global Nash equilibrium by providing an online privacy information encrypted learning algorithm, where the encryption-decryption is embedded in the data transmission and computation to overcome the potential attacks in unreliable networks. During the encrypted learning process, the global Nash equilibrium is decentralized to the sum of a set of local performance indices, and the pair of optimal gains is decoupled to integrate the global solution.
- The convergence of the decentralized learning algorithm for global Nash equilibrium is achieved successfully, then the developed encrypted learning process is implemented with a novel quantization scheme and an additive-multiplicative property. It is an intractable issue to solve the performance index of global Nash equilibrium as the monotonicity can hardly be obtained in the $\text{maxmin}$ problem.

### 2 Preliminaries

Let $\mathcal{V} = \{1, 2, \ldots, N, N + 1\}$ be the set of $N + 1$ agents with $i \in \mathcal{V}$ representing the $i$th agent. A weighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is used to characterize the interaction among agents, where $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ denotes the edge set with paired agents. An edge $(j, i) \in \mathcal{E}$ means that the $i$th agent can receive information from the $j$th agent. The neighborhood set $\mathcal{N}_i$ of agent $i$ is defined as $\mathcal{N}_i = \{j | (j, i) \in \mathcal{E}\}$. The adjacency matrix is defined as $A_{\mathcal{G}} = [a_{ij}]_{N+1 \times N+1}$, where $a_{ii} = 0$, $a_{ij} > 0$ if $(j, i) \in \mathcal{E}$, and $a_{ij} = 0$, otherwise. The graph is
undirected if \( a_{ij} = a_{ji} \) for all \((j, i) \in E\), else it is directed. The Laplacian matrix \( L = \{L_{ij}\}_{N \times N} \) is defined as \( L_{ii} = \sum_{j \in N_i} a_{ij}, \) \( L_{ij} = -a_{ij} \) for \( i \neq j \), which is also equal to \( L = D - A_{\text{adj}}, \) \( D = \text{diag}(d_i) \) is called the in-degree matrix, where \( d_i = \sum_{j \in N_i} a_{ij} \). A directed path on \( G \) from agent \( i_1 \) to agent \( i_2 \) is a sequence of ordered edges in the form of \((i_k, i_{k+1}) \in E, k = 1, 2, \ldots, l - 1\). A graph contains a directed spanning tree \( T \) if it has at least one agent with directed paths to all other agents.

The hierarchical structure of human-cyber-physical control systems is considered (Zhu and Basar, 2015), where cyber layer directs the physical layer to implement some desired actions or policies, and human layer manages the information process. To guarantee the security of information among the agents in cyber and physical layers, as presented in Fig. 1, the semi-homomorphic encryption technique is embedded in the cloud-based parameter computation between physical and cyber layers by human layer.

![Physical layer](Image 372 to 463)  
Physical layer
![Cyber layer](Image 376 to 390)  
Cyber layer
![Human layer](Image 381 to 385)  
Human layer

Suppose that in the multi-agent system, leader may send information to the followers, but does not receive information from any one of the followers. It means that with the leader being the root node, there exists a directed path from the leader to any one of the followers. Let the state of the leader node be \( x_1(t) \in \mathbb{R}^n \), and assumed to satisfy the dynamic as:

\[
x_1(t + 1) = Ax_1(t).
\]

For the followers, the discrete-time dynamic is determined by the following form

\[
x_i(t + 1) = Ax_i(t) + B\gamma(t)u_i(t) + Dw_i(t),
\]

where \( x_i(t) \in \mathbb{R}^n \) is the system state, \( \gamma(t) \in \{0, 1\} \) denotes the stochastic variable of packet loss information in communication channel, \( u_i(t) \in \mathbb{R}^{m_1} \) and \( w_i(t) \in \mathbb{R}^{m_2} \) are the control input and the disturbance input in node \( i, i = 2, \ldots, N + 1 \).

Define the information structure of such a multi-agent system as

\[
\begin{align*}
\mathcal{F}(0) &:= \{x_i(0), i = 1, 2, \ldots, N + 1\}, & t &= 0 \\
\mathcal{F}(t) &:= \{\Xi(t), \Gamma(t - 1)\}, & t &= 1, \ldots
\end{align*}
\]

where \( \Xi(t) = (x_i(1), x_i(2), \ldots, x_i(t); i = 1, 2, \ldots, N + 1) \), and \( \Gamma(t - 1) = (\gamma(0), \gamma(1), \ldots, \gamma(t - 1)) \). In the leader-follower consensus problem, the Laplacian matrix is

\[
L = \begin{bmatrix}
0 & 0 & \cdots & 0 \\
-\alpha & L
\end{bmatrix}
\]

where \( 0_N = [0, 0, \ldots, 0]^T, \) \( \alpha = [a_{21}, a_{31}, \ldots, a_{N+1,1}]^T, \) and

\[
L = \begin{bmatrix}
\sum_{j \in N_2} a_{2j} & -a_{23} & \cdots & -a_{2, N + 1} \\
-a_{32} & \sum_{j \in N_2} a_{3j} & \cdots & -a_{3, N + 1} \\
\vdots & \vdots & \ddots & \vdots \\
-a_{N, N+1, 2} & -a_{N, N+1, 3} & \cdots & \sum_{j \in N_{N+1}} a_{N, N+1, j}
\end{bmatrix}
\]

**Assumption 1** All the eigenvalues of \( A \) are either on or outside the unit disk.

**Remark 1** Assumption 1 makes the open loop system of leader unstable. Based on the circle criterion, it is an effective assumption to avoid the trivial case that all the states of followers can converge to zero without consensus protocols.

**Assumption 2** The stochastic variable sequence \( \{\gamma(t)\}_{t \geq 0} \) is distributed as an i.i.d. Bernoulli process with \( \mathbb{P}(\gamma(t) = 1) = \mu. \)

**Remark 2** This work focuses on the consensus problem under potential attackers, where the packet losses are caused by a malicious jammer and can be considered identical in the system (Xu et al., 2016, 2019). Besides, the followers’ dynamics (2) with packet loss information (3) can model the TCP-like network protocol with denial-of-service attacks (Moon and Bañar, 2014; Zhu and Basar, 2015).

It can be easily obtained that mean \( \mathbb{E}(\gamma(t)) = \mu \) and variance \( \mathbb{D}(\gamma(t)) = \mathbb{E}(\gamma^2(t)) - (\mathbb{E}(\gamma(t)))^2 = \mu - \mu^2 . \) The interaction among agents is characterized by a connected graph \( G = (V, E) \). For a soft-constrained dynamic game, the consensus and disturbance policies for followers are considered by

\[
\begin{align*}
u_i(t) &= K_1[\sum_{j \in N_i} a_{ij}(x_j(t) - x_i(t))], \\
w_i(t) &= -K_2[\sum_{j \in N_i} a_{ij}(x_j(t) - x_i(t))]
\end{align*}
\]

where the disturbance policy \( w_i(t) \) is regarded as a policy against consensus policy \( u_i(t) \), which aims for the worst consensus performance.
Augment state $\bar{x}(t) = [x_1(t)^T, x_2(t)^T, \cdots, x_{N+1}(t)^T]^T$, then, from the dynamics (1) and (2), the system can be rewritten as

$$\bar{x}(t+1) = (I_{N+1} \otimes A + \gamma(t)\mathcal{L} \otimes B K_1 - \mathcal{L} \otimes D K_2)\bar{x}(t)$$

where $\otimes$ means Kronecker product.

Let $\delta(t) = [(I_{N+1} - M_{N+1}) \otimes I_n]\bar{x}(t)$, where $M_{N+1} = [\mathcal{M}_1, \mathcal{M}_2]^T$, $\mathcal{M}_1 = 0_{N+1}$ and $\mathcal{M}_2 = [I_N, 0_{N,N}]^T$. It is easy to know that 1 is the eigenvalue of $(I_{N+1} - M_{N+1})$ with multiplicity $N + 1$. Thus, the consensus error dynamics become

$$\delta(t + 1) = [(I_{N+1} - M_{N+1}) \otimes I_n] \bar{x}(t + 1)$$

$$= (I_{N+1} \otimes A + \gamma(t)\mathcal{L} \otimes B K_1 - \mathcal{L} \otimes D K_2)\delta(t).$$

Consider $\bar{\delta}(t) = ([0_{N-1}, I_N] \otimes I_n)\delta(t) = [\delta_2(t), \ldots, \delta_{N+1}(t)]^T$, one has

$$\bar{\delta}(t + 1) = (I_N \otimes A + \gamma(t)\mathcal{L} \otimes B K_1 - \mathcal{L} \otimes D K_2)\bar{\delta}(t),$$

and it follows that $\bar{\delta}(t) = 0$ if and only if $x_1(t) = x_2(t) = \cdots = x_{N+1}(t)$.

**Definition 1 (Aliyu, 2011)** The system (8) is said to have $L_2$-gain less than or equal to $\eta$, if for any initial state $\bar{\delta}(0)$ and $\bar{x}(0)$, the response $\bar{z}(t) = C(t)\bar{\delta}(t)$ corresponding to any $\bar{x}(t) \in L_2[0, \infty)$ satisfies:

$$\mathbb{E}\{\sum_{t=0}^{\infty} ||z(t)||^2 | F(0)\} \leq \eta^2 \sum_{t=0}^{\infty} ||\bar{x}(t)||^2 + \phi(\delta(0))$$

for some bounded function $\phi$ such that $\phi(0) = 0$.

**Bounded $L_2$-gain consensus problem:** (Zhou et al., 1996) For the multi-agent system (8), it is desired to find the control input $u_i(t)$ to solve the standard consensus problem when disturbance input $w_i(t) = 0$; and also satisfy the $L_2$-gain condition (9) for a given $\eta$ when $w_i(t) \neq 0$, with initial state $\delta(0) = 0$.

To guarantee the condition (9), the soft-constrained dynamic game can be utilized to construct the $H_\infty$ optimal control (Başar and Bernhard, 2008). We define that $C^T(t)C(t) = [Q_N \otimes I_n + \gamma(t)(L \otimes B K_1)^T S (L \otimes B K_1)]$, $S = I_N \otimes S$, $Q_N \in \mathbb{R}^{N \times N}$ and $S \in \mathbb{R}^{m_1 \times m_1}$ are symmetrical positive definite matrices, $\gamma \geq \gamma^*$, and $\gamma^*$ is the smallest $L_2$-gain value of the system.

For the cyber layer in the hierarchical structure, privacy violation is also a critical security threat from network attackers, rather than merely packet losses. To achieve the secure and privacy-preserving consensus, the encryption-decryption is embedded in the data transmission and computation, and an encrypted learning scheme is developed with the encryption technique. Next we give the following encryption-decryption property.

**Definition 2 (Marquia et al., 2020)** Assume there exist operators $\circ$ and $\circ$ such that $(\mathbb{P}, \circ)$ and $(\mathbb{C}, \circ)$ form groups with plaintext set $\mathbb{P}$ and ciphertext set $\mathbb{C}$. A public key encryption $(\mathbb{P}, \mathbb{C}, K, \mathcal{E}, \mathcal{D})$ is called homomorphic if $\mathcal{D}(\mathcal{E}(x_1, \kappa_p) \circ \mathcal{E}(x_2, \kappa_p), \kappa_p, \kappa_s) = x_1 \times x_2$ for all $x_1, x_2 \in \mathbb{P}$ and $\kappa_p, \kappa_s \in \mathbb{K}$, where $\mathbb{K}$ denote key sets, $\mathcal{E}$ and $\mathcal{D}$ are encryption and decryption operations.

**Paillier’s semi-homomorphic encryption:** (Hadjicostis and Domingo-Ferrer, 2020; Ruan et al., 2019) The encryption method in this paper is used with the following schemes: (1) Randomly select two large prime numbers $a_1$ and $a_2$ such that $\gcd(a_1 a_2, (1 - a_1)(1 - a_2)) = 1$, $\gcd(1, \cdot)$ means the greatest common divisor; (2) Use the public key $\kappa_p = a_1 a_2$, private key $\kappa_s = (b_1, b_2)$ with $b_1 = \phi(\kappa_p) = \text{lcm}(a_1 - 1, a_2 - 1)$, $b_2 = \phi^{-1}(\kappa_p)$ mod $\kappa_p$, where lcm means the least common multiple, $\phi(\cdot)$ is Euler’s totient function, and $\phi^{-1}(\cdot)$ is the modular multiplicative inverse of $\phi(\cdot)$; (3) The positive integer plaintext $c_1 \in \mathbb{Z}_{\kappa_p} = \mathbb{P}$ is encrypted by $\mathcal{E}(c_1, \kappa_p) = \gamma^{c_1} \cdot d^{a_2} \text{mod} \kappa_p$, $d$ is a random number; (4) The ciphertext $c_2 \in \mathbb{Z}_{\kappa_p}^2 = \mathbb{C}$ is decrypted by $\mathcal{D}(c_2, \kappa_p, \kappa_s)$.

**3 Global Nash equilibrium with a privacy information encrypted learning algorithm**

Generally, for the leader-follower consensus process with $N + 1$ agents, the global Nash Equilibrium is difficult to achieve. Due to the large size of the system, the coupled dynamics and policies complicate the game’s solution, even if all the state information is available. To overcome the critical issue, we first decentralize the global Nash equilibrium into a set of local performance indices, find the necessary and sufficient condition of its existence, then decouple the policies to integrate the privacy information encrypted learning structure.

### 3.1 The decentralized zero-sum game

Define the global performance index for dynamic (8) as

$$J_v(\delta(t), K_1, K_2) = \mathbb{E}\{\sum_{k=1}^{\infty} \delta(k)^T Q_{(K_1, K_2)}(t) \delta(k) | F(t)\}$$

(10)
where \( Q_{(K_1, K_2)}(t) = Q_N \otimes I_n + \gamma(t)(L \otimes BK_1)^T S(L \otimes BK_1) - \eta^2 (L \otimes DK_2)^T (L \otimes DK_2) \), \( \gamma(t) \) denotes the stochastic variable, \( Q_N = \text{diag}(\lambda_1^2, \ldots, \lambda_N^2) \), and \( \lambda_i \) is the \( i \)th eigenvalue of matrix \( L \). To make sure that matrix \( Q_N \) is positive definite as in condition (9), all eigenvalues of \( L \) should be nonzero, for simplicity, the communication between the followers is considered as undirected throughout this paper.

Based on the information structure (3) and optimal control theory (Anderson and Moore, 1971; Levis et al., 1971), we can rewrite the steady-state performance index (10) in an infinite horizon as \( J_o(\delta(t), K_1, K_2) \triangleq \delta(t)^T \mathcal{P} \delta(t) \), \( \mathcal{P} = \mathcal{P}^T > 0 \), and further obtain

\[
J_o(\delta(t), K_1, K_2) = \mathbb{E}\{\delta(t)^T Q_{(K_1, K_2)}(t) \delta(t) | \mathcal{F}(t)\} \\
+ \mathbb{E}\{J_o(\delta(t+1), K_1, K_2) | \mathcal{F}(t)\}
\]

\( (11) \)

with the system dynamic (8).

Then the global Nash equilibrium solution \((K_1^*, K_2^*)\) of thus a zero-sum game satisfies

\[
J_o(\delta(0), K_1^*, K_2^*) = \min_{K_1} \max_{K_2} J_o(\delta(0), K_1, K_2)
\]

\( (12) \)

where policies \( u_i(t) \) and \( w_i(t) \) are determined by the control gains \( K_1 \) and \( K_2 \), respectively. In the soft-constrained zero-sum game, the control policy \( u_i(t) \) seeks to minimize index (10), and the disturbance policy \( w_i(t) \) seeks to maximize the index.

**Theorem 1** The multi-agent system (8) satisfies the condition (9), if and only if the global Nash equilibrium (12) exists and is unique.

**Proof.** (Necessary) Suppose that the global Nash equilibrium (12) exists and is unique, the optimal gains are \( K_1^* \) and \( K_2^* \), then the optimal performance index is

\[
J_o(\delta(t), K_1^*, K_2^*) = \mathbb{E}\{\delta(t)^T Q_{(K_1^*, K_2^*)} (t) \delta(t)\} + J_o(\delta(t+1), K_1^*, K_2^*) | \mathcal{F}(t)\}
\]

\( (13) \)

which gives

\[
\max_{K_2} \left\{ \mathbb{E}\{J_o(\delta(t+1), K_1^*, K_2^*) - J_o(\delta(t), K_1^*, K_2^*) \}
+ \delta(t)^T Q_{(K_1^*, K_2^*)} (t) \delta(t) | \mathcal{F}(t)\} \right\} = 0.
\]

\( (14) \)

For any other \( K_2 \neq K_2^* \) and time \( t_s > 0 \), it yields

\[
\mathbb{E}\{J_o(\delta(t_s+1), K_1^*, K_2^*) - J_o(\delta(0), K_1^*, K_2^*) | \mathcal{F}(0)\}
+ \mathbb{E}\{\sum_{t=0}^{t_s} \delta(t)^T Q_{(K_1^*, K_2^*)} (t) \delta(t) | \mathcal{F}(0)\} \leq 0,
\]

\( (15) \)

and taking \( t_s \to \infty \) with initial state \( \delta(0) = 0 \), we have

\[
\mathbb{E}\{\sum_{t=0}^{\infty} ||z(t)||^2 | \mathcal{F}(0)\} \leq \eta^2 \sum_{t=0}^{\infty} ||z(t)||^2.
\]

\( (16) \)

(Sufficient) Considering the condition (9) holds, the global performance index (10) is bounded with admissible controls as \( J_o(\delta(t), K_1, K_2) < \infty \). The function \( J_o(\delta(t), K_1, K_2) \) is strictly convex with respect to (w.r.t) \((I \otimes K_1) \delta(t)\) and strictly concave w.r.t \((I \otimes K_2) \delta(t)\). Based on the Theorem 2.3 by Ba¸sar and Bernhard (2008), the performance index (10) will produce a unique Nash equilibrium (12). The proof is thus completed. \( \square \)

The global performance index (11) consists of \( N+1 \) interconnected agents’ dynamics. To analyze the solution of global Nash equilibrium, the scheme is detailed in the follows.

**Lemma 1** The performance index of global Nash equilibrium (12) can be decentralized to the sum of a set of local performance indices, if we define the local performance index for node \( i \) as

\[
J_i(\xi_i(t), K_1, K_2) = \mathbb{E}\{\sum_{k=i}^{\infty} \xi_i(k)^T \bar{Q}(t) \xi_i(k) | \mathcal{F}(t)\}
\]

\( (17) \)

for the local dynamic

\[
\xi_i(t+1) = (A + \gamma(t) \lambda_i B K_1 - \lambda_i D K_2) \xi_i(t)
\]

\( (18) \)

where \( \bar{Q}(t) = I + \gamma(t)(B K_1)^T S(B K_1) - \eta^2 (D K_2)^T (D K_2) \). \( \bar{Q}(t) \)

**Proof.** Consider a pair \((K_1^*, K_2^*)\) being the solution of global Nash equilibrium, which holds for the following optimization problem

\[
f_0 : \min_{K_1, K_2} \max_{K_1, K_2} J_o(\delta(t), K_1, K_2)
\]

s.t. \( \delta(t+1) = (I_N \otimes A + \gamma(t) L \otimes B K_1 - L \otimes D K_2) \delta(t) \).

\( (19) \)

Let \( U \in \mathbb{R}^{N \times N} \) be such a unitary matrix that \( U^T U = I_N \), and \( U L U^T = L \triangleq \text{diag}(\lambda_1, \ldots, \lambda_N) \). By choosing \( \xi(t) = [\xi_1(t), \ldots, \xi_N(t)] = (U \otimes I_n) \delta(t) \), it is not difficult
to obtain that
\[
J_{\circ}(\bar{\delta}(t), K_1, K_2) = \mathbb{E}\{\sum_{k=1}^{\infty} \xi(k)^T (U \otimes I_n) [Q_N \otimes I_n + \gamma(t) \\
\times (L \otimes BK_1)^T S (L \otimes BK_1) - \eta^2 (L \otimes DK_2)^T \}
\]
\[
\times (L \otimes DK_2) [(U \otimes I_n)^T \xi(k) | F(t)]\}
\]
\[
= \mathbb{E}\{\sum_{k=1}^{N} \xi_i(k)^T [\lambda_i^2 I_n + \gamma(t) \lambda_i^2 (BK_1)^T S \xi_i(t)] \\
\times (BK_1) - \eta^2 \lambda_i^2 (DK_2)^T (DK_2) \xi_i(t) | F(t)]\}
\]
\[
= \sum_{i=1}^{N} \lambda_i^2 J_i(\xi_i(t), K_1, K_2).
\]
For the dynamic (8), it becomes
\[
\xi_i(t + 1) = (I_N \otimes A + \gamma(t) \Lambda \otimes BK_1 - \Lambda \otimes DK_2) \xi_i(t),
\]
which is equivalent to the simultaneous dynamic
\[
\xi_i(t + 1) = (A + \gamma(t) \lambda_i BK_1 - \lambda_i DK_2) \xi_i(t)
\]
for all \(i = 1, 2, \ldots, N\).

Based on the transform by \(\bar{\delta}(t) = (U \otimes I_n)^T \xi(t)\), the problem (19) is equivalent to the optimization
\[
f_1 : \min_{K_1} \max_{K_2} \sum_{i=1}^{N} \lambda_i^2 J_i(\xi_i(t), K_1, K_2)
\]
s.t. \(\xi_i(t + 1) = (A + \gamma(t) \lambda_i BK_1 - \lambda_i DK_2) \xi_i(t)\)
for all \(i = 1, 2, \ldots, N\). This completes the proof. \(\square\)

From the Lemma 1, the pair \((K_1^*, K_2^*)\) of global Nash equilibrium is equal to the solution of optimization (23). According to the steady-state performance index in an infinite horizon (Levis et al., 1971), for the local performance index (17) w.r.t the local dynamic (18), there exists a matrix \(P_i = P_i^T > 0\) such that \(J_i(\xi_i(t), K_1, K_2) = \xi_i^T(t) P_i \xi_i(t)\).

**Theorem 2** The global Nash equilibrium exist and is unique if and only if (a) the condition \(\mathcal{H} > 0\); or (b) the pair of optimal gains of global Nash equilibrium satisfies

\[
\begin{cases}
K_1^* = -\mathcal{Y}^{-1} \sum_{i=1}^{N} \lambda_i B^T P_i (A - \lambda_i D K_2^*) \\
K_2^* = -\mathcal{H}^{-1} \sum_{i=1}^{N} \lambda_i D^T P_i (A + \mu \lambda_i B K_1^*)
\end{cases}
\]

where \(\mathcal{Y} = \sum_{i=1}^{N} F_i, F_i = B^T S B + \lambda_i^2 B^T P_i B, \mathcal{H} = \sum_{i=1}^{N} H_i, H_i = (\eta^2 D^T D - \lambda_i^2 D^T P_i D)\).

**Proof.** (Necessary) Based on Lemma 1, the global Nash equilibrium can be decentralized to the optimization (23). Considering the decentralized performance index for any \(i\), we have

\[
J_i(\xi_i(t), K_1, K_2) = \mathbb{E}\{\xi_i^T(t) \tilde{Q}(t) \xi_i(t) | F(t)\}
\]
\[
+ \mathbb{E}\{J_i(\xi_i(t+1), K_1, K_2) | F(t)\}
\]
where

\[
\mathbb{E}\{J_i(\xi_i(t+1), K_1, K_2) | F(t)\} = \mathbb{E}\{\xi_i^T(t + 1) \tilde{Q}(t + 1) \xi_i(t + 1) | F(t)\}
\]
\[
= \mathbb{E}\{\xi_i^T(t + 1) (I - \eta^2 (DK_2)^T (DK_2)) \xi_i(t + 1) | F(t)\}
\]
\[
= J_i(\xi_i(t + 1), K_1, K_2) | F(t)\}
\]

By inserting (26) into (25), we have:

\[
\xi_i^T(t) P_i \xi_i(t) = \xi_i^T(t) [I - \eta^2 (DK_2)^T (DK_2)]
\]
\[
+ (A + \mu \lambda_i BK_1 - \lambda_i DK_2)^T P_i (A + \mu \lambda_i BK_1 - \lambda_i DK_2) \]
\[
- \lambda_i DK_2) + \mu (BK_1)^T [S + (1 - \mu) \lambda_i^2 P_i]
\]
\[
\times (BK_1) | F(t)\}
\]

Take the derivative of (27) w.r.t control input \(u_i(t) = K_i \xi_i(t)\), we can obtain the optimal control gain satisfies:

\[
F_i K_i^* = -\lambda_i B^T P_i (A - \lambda_i D K_2^*).
\]

Summing equation (28) with \(i = 1, 2, \ldots, N\), it becomes

\[
\sum_{i=1}^{N} F_i K_i^* = -\sum_{i=1}^{N} \lambda_i B^T P_i (A - \lambda_i D K_2^*),
\]
which yields the optimal control gain

\[
K_1^* = -\mathcal{Y}^{-1} \sum_{i=1}^{N} \lambda_i B^T P_i (A - \lambda_i D K_2^*).
\]

By using the similar way, derivative (27) w.r.t. disturbance input \(w_i(t) = K_2 \xi_i(t)\) and sum for all terms, we can obtain the optimal disturbance gain satisfies:

\[
\sum_{i=1}^{N} H_i K_2^* = -\sum_{i=1}^{N} \lambda_i D^T P_i (A + \mu \lambda_i B K_1^*).
\]
When the local Nash equilibrium exist and is unique, it must require that the game is strictly concave in $w_i(t) = K_2 \xi_i(t)$ as $\mathcal{H} > 0$, which means

\begin{equation}
K_2^* = -\mathcal{H}^{-1} \sum_{i=1}^{N} \lambda_i D^T P_i (A + \mu \lambda_i B K_1^*). \tag{32}
\end{equation}

(Sufficient) Suppose that the condition (a) $\mathcal{H} > 0$ holds or (b) optimal gain solution satisfies (24), when inserting the optimal gains $K_1^*$ and $K_2^*$ into the local performance index, the following equation holds as

\begin{equation}
(K_1^*, K_2^*) = \arg \min_{K_1, K_2} \sum_{i=1}^{N} \lambda_i^2 J_i(\xi_i(t), K_1, K_2) \tag{33}
\end{equation}

\begin{equation}
= \arg \min_{K_1, K_2} J_o(\delta(t), K_1, K_2)
\end{equation}

which indicates that the pair $(K_1^*, K_2^*)$ will arrive the global Nash equilibrium. It completes the proof. \hfill \Box

**Remark 3** From Lemma 1, the global Nash equilibrium is a necessary but possibly insufficient condition to the simultaneous local Nash equilibriums with the following optimization

\begin{equation}
\min_{K_1, K_2} J_o(\delta(t), K_1, K_2) \tag{34}
\end{equation}

s.t. $\xi_i(t + 1) = (A + \gamma(t) \lambda_i B K_1 - \lambda_i D K_2) \xi_i(t)$

for each $i$. It can also be easily obtained as $H_i > 0 \Rightarrow \mathcal{H} > 0$ and $H_i > 0 \Leftrightarrow \mathcal{H} > 0$ based on Theorem 2. Different with the typical distributed zero-sum game (Jiao et al., 2016), we don’t require that the local Nash equilibriums hold simultaneously.

In this subsection, the global Nash equilibrium is successfully decentralized into the sum of a set of local performance indices, which yields a pair of coupled optimal control gains (24). The necessary and sufficient condition of existence of global Nash equilibrium is discussed in Theorem 2, and the decentralized equation (27) actually provides a medium for solving the global Nash equilibrium in the following algorithm.

### 3.2 The online privacy information encrypted learning algorithm

To further solve the global Nash equilibrium from the set of decentralized local performance indices, we decouple the pair of gains (24) in the following lemma.

**Lemma 2** The coupled control gains with the form (24) for the global Nash equilibrium can be decoupled as

\begin{equation}
\begin{aligned}
K_1^* &= -M_1^{-1} \mathcal{H}^{-1} \left( \sum_{i=1}^{N} \lambda_i B^T P_i A \right) \\
&\quad + \sum_{i=1}^{N} \lambda_i^2 B^T P_i D \mathcal{H}^{-1} \left( \sum_{i=1}^{N} \lambda_i D^T P_i A \right); \\
K_2^* &= -M_2^{-1} \mathcal{H}^{-1} \left( \sum_{i=1}^{N} \lambda_i D^T P_i A \right) \\
&\quad - \sum_{i=1}^{N} \mu \lambda_i^2 D^T P_i B \mathcal{H}^{-1} \left( \sum_{i=1}^{N} \lambda_i B^T P_i A \right)
\end{aligned}
\end{equation}

where $M_1 = I_{m_1} + \mathcal{H}^{-1} \sum_{i=1}^{N} \lambda_i^2 B^T P_i D \mathcal{H}^{-1} \sum_{i=1}^{N} \mu \lambda_i^2 D^T P_i B$, $M_2 = I_{m_2} + \mathcal{H}^{-1} \sum_{i=1}^{N} \mu \lambda_i^2 D^T P_i B \mathcal{H}^{-1} \sum_{i=1}^{N} \lambda_i^2 B^T P_i D$, if the global Nash equilibrium exists.

**Proof.** As matrices $M_1$ and $M_2$ are invertible by definitions, the gains can be easily decoupled to (35) by using equation (24), when the global Nash equilibrium exists. It gives the proof. \hfill \Box

Combining Theorem 2, the decoupled optimal gain pair (35) satisfies the set of decentralized local performance indices (27) simultaneously, which can provide an iterative way to solve the global Nash equilibrium. As presented in Algorithm 1, some iterative functions are listed as follows:

\begin{align*}
f_2(P_i^{(l)}) &= I_{m_1} + \mathcal{H}^{(l)} \mathcal{H}^{-1} \sum_{i=1}^{N} \mu \lambda_i^2 B^T P_i^{(l)} D \mathcal{H}^{-1} \sum_{i=1}^{N} \lambda_i D^T P_i^{(l)} B; \\
f_3(P_i^{(l)}) &= I_{m_2} + \mathcal{H}^{(l)} \mathcal{H}^{-1} \sum_{i=1}^{N} \mu \lambda_i^2 D^T P_i^{(l)} B \mathcal{H}^{-1} \sum_{i=1}^{N} \lambda_i^2 B^T P_i^{(l)} D; \\
f_4(P_i^{(l)}) &= -M_1^{(l)} \mathcal{H}^{(l)} \mathcal{H}^{-1} \sum_{i=1}^{N} \lambda_i B^T P_i^{(l)} A \\
&\quad + \sum_{i=1}^{N} \lambda_i^2 B^T P_i^{(l)} D \mathcal{H}^{(l)} \mathcal{H}^{-1} \sum_{i=1}^{N} \lambda_i D^T P_i^{(l)} A; \\
f_5(P_i^{(l)}) &= -M_2^{(l)} \mathcal{H}^{(l)} \mathcal{H}^{-1} \sum_{i=1}^{N} \lambda_i D^T P_i^{(l)} A \\
&\quad - \sum_{i=1}^{N} \mu \lambda_i^2 D^T P_i^{(l)} B \mathcal{H}^{(l)} \mathcal{H}^{-1} \sum_{i=1}^{N} \lambda_i B^T P_i^{(l)} A)
\end{align*}

where $\mathcal{H}^{(l)} = \sum_{i=1}^{N} F_i^{(l)}$ and $\mathcal{H}^{(l)} = \sum_{i=1}^{N} H_i^{(l)}$. 

\hfill \Box

7
Algorithm 1 Unprotected Decentralized Iterative Learning Algorithm for Global Nash equilibrium

Initialization

Select initial admissible gain pair \((K_1^{(1)}, K_2^{(1)})\) and a constant \(\epsilon > 0\) small enough, and set \(l = 1\).

Procedural

1: \textbf{while} \(\max_i \| P_i^{(l)} - P_i^{(l-1)} \| \geq \epsilon\), for \(l > 1\) \textbf{do}
2: \hspace{1em} \textbf{for} \(i = 1, \ldots, N\) \textbf{do}
3: \hspace{2em} Solve local matrix \(P_i^{(l)}\) from equation (27)
4: \hspace{2em} Tune the parameters as
5: \hspace{3em} \(F_i^{(l)} = [B^T S B + \lambda_2^2 B^T P_i^{(l)} B]\); \hspace{1em} \(H_i^{(l)} = [\eta^2 D^T D - \lambda_2^2 D^T P_i^{(l)} D]\);
6: \hspace{1em} \textbf{end for}
7: \hspace{1em} Learn the matrices with
8: \hspace{2em} \(M_1^{(l)} \leftarrow f_2(P_i^{(l)})\); \(M_2^{(l)} \leftarrow f_3(P_i^{(l)})\)
9: \hspace{1em} \textbf{end while}
10: \textbf{End Procedure}

Theorem 3 The decentralized iterative learning algorithm will produce a sequence of gain pairs \(\{(K_1^{(l)}, K_2^{(l)})\}, \ l = 1, 2, \ldots\), which converges to the optimal gain pair \((K_1^*, K_2^*)\) as \(l \to \infty\), if the global Nash equilibrium exists.

Proof. The proof is given in Appendix. \(\Box\)

In Algorithm 1, the set of decentralized local performance indices (27) and the decoupled pair of gains (35) provide an effective way to solve the global Nash equilibrium for the leader-follower system. Theorem 3 gives the convergence with an iterative learning process successfully, where as the competitive relationship of policies in a zero-sum game, the convergence of \textit{maxmin} solution is generally hard to be obtained.

However, the procedure in Algorithm 1 involves lots of unprotected privacy data in the networked system. To further guarantee the privacy information of states and controls secure under the hierarchical structure, the data transmissions and operations should be encrypted between cyber layer and physical layer, including the iterative computations of decentralized equation (27) in the cloud servers. The implementation of a developed encrypted learning process is detailed in the follows.

Based on equation (27), we have

\[
\begin{align*}
\xi_i^T(t)P_i^l(t) &= \xi_i^T(t)[I + \mu(BK_1)^T S(BK_1) - \eta^2(DK_2)^T] \\
&\times (DK_2)[\xi_i(t) + \xi_i^T(t+1)P_i^l(t+1)]
\end{align*}
\]

which can be vectored as

\[
\begin{align*}
\tilde{Z}_i(t)^T &\quad \text{balance digits} \\
\tilde{Y}_i(t) &\quad \text{quantize} \\
\tilde{Y}_i(t) &\quad \text{reset by} \quad N_i \\
\tilde{Z}_i(t) &\quad \text{learn} \\
\end{align*}
\]

\[
\begin{align*}
\tilde{Z}_i(t)^T &= \left[ \frac{\xi_i(t)\xi_i^T(t)}{\xi_i(t+1)\xi_i^T(t+1)} \right] \tilde{P}_i = \\
&\quad \frac{\xi_i^T(t)[I + \mu(BK_1)^T S(BK_1) - \eta^2(DK_2)^T] \xi_i(t)}{\xi_i(t+1)\xi_i^T(t+1)}
\end{align*}
\]

where \(\tilde{P}_i \in \mathbb{R}^n\), \(\bar{n} = \frac{n(n+1)}{2}\) as it is symmetrical.

Note that in the equation (40), vector \(\tilde{P}_i\) is unknown and need to be solved, and also privacy as it provides the core control parameters for achieving global Nash equilibrium. \(Y_i(t)\) is a measured utility value in practice, and \(\tilde{Z}_i(t)\) is a vector consisting of consensus states directly, which are also sensitive data in privacy preserving. With enough sampling data, the equation (40) can be further augmented as \(\tilde{P}_i = \tilde{Z}_i^TY_i\), where \(Y_i = [Y_i(t_1), \ldots, Y_i(t_n)]^T\), \(Z_i = [\tilde{Z}_i(t_1), \ldots, \tilde{Z}_i(t_n)]^{-1}\) if \([\tilde{Z}_i(t_1), \ldots, \tilde{Z}_i(t_n)]\) is an invertible matrix.

To make sure the information security of states and controls, matrix \(Z_i\) should be privacy in the cloud-based computation, thus we use the Paillier’s encryption method to encrypt the data before transmissions and operations. The detailed privacy data flow in the encrypted learning architecture is presented in Fig. 2.

Fig. 2. The detailed privacy data flow in the encrypted learning architecture

Paillier’s encryption scheme can only work with finite ring of positive integers. For this reason, a balanced quantization scheme is first proposed here to detail the encrypted learning process.

For any parameter \(\rho \in \mathbb{R}\), there are always \(r_1, r_2 \in \mathbb{N}\) such that

\[
\frac{r_1}{r_2} \leq \rho \leq \frac{r_1 + 1}{r_2}.
\]

A common multiple \(r_m \in \mathbb{N}\) can be found for all the parameters to be quantized, then, we can quantize the parameter \(\rho\) by \(q_1(\rho) = \lfloor r_m\rho \rfloor\) with a max quantization
error $1/r_m$, which can be guaranteed small enough by selecting a large $r_m$. Besides, to make a balanced quantization, we reduce the difference of parameter digits in matrices $Z_i$ and $Y_i$ by selecting a balance factor $v > 0$ such that $vZ_i$ and $Y_i/v$ are in some similar digits.

**Remark 4** As pointed out in (Kogiso and Fujita, 2015), the errors caused by quantization is still open in changing performance of the networked control system. Different from the common fixed-point arithmetic (Ruan et al., 2019; Murguia et al., 2020), the balance factor $v$ designed here can mitigate the influence of quantization errors in the data transmissions and computing processes.

Thus, the matrices $Z_i$ and $Y_i$ can be quantized to integer parameter matrices $\tilde{Z}_i$ and $\tilde{Y}_i$ by using balance factor $v$ and $r_m$, such that $\tilde{P}_i = (vZ_i)^T \cdot (Y_i/v) = \tilde{Z}_i^T \cdot \tilde{Y}_i + \varepsilon(1/r_m)$, where $\|\varepsilon(1/r_m)\|$ can be arbitrary small, even $\|\varepsilon(1/r_m)\| = 0$ if $q_i(\rho) = r_m \rho$, for all parameter $\rho$ in matrices $vZ_i$ and $Y_i/v$. Moreover, let a parameter $\rho^* \in \tilde{P}_i$, which can be inverse quantized as $q_2(\rho^*) = \rho^*/r_m^2$.

To further make all the computing parameters be positive integers, we reset parameter $\theta \in Z_i \cup Y_i$ by

$$\hat{\theta} = \begin{cases} \theta + N \theta, & \theta < 0 \\ \theta, & \theta \geq 0 \end{cases}$$

where $\tilde{Z}_i \cup \tilde{Y}_i$ denotes the set of all parameters in $\tilde{Z}_i$ and $\tilde{Y}_i$, $N_\theta = 2 \theta_m + 1$ is the resetting factor, and $\theta_m = \max(\theta, \theta \in \text{abs}(Z_i \cup Y_i))$. Then, matrices $\tilde{Z}_i$ and $\tilde{Y}_i$ can be reset to the positive integer parameter matrices $\tilde{Z}_i$ and $\tilde{Y}_i$ by using (42). Denote $[\cdot]_{kh}$ as the matrix element of row $k$ and column $h$, the online encrypted learning process is developed in Algorithm 2.

For the encrypted computation, we give the following additive-multiplicative property.

**Lemma 3** For any parameters $\theta_1, \theta_2, \theta_3, \theta_4 \in \tilde{Z}_i \cup \tilde{Y}_i$, the corresponding positive integer parameters $\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \hat{\theta}_4$ in $\tilde{Z}_i \cup \tilde{Y}_i$ are reset as equation (42), then, by using Paillier’s semi-homomorphic encryption, we can obtain the following additive-multiplicative property:

$$(\theta_1 \theta_2 + \theta_3 \theta_4) = \mathcal{D}((\mathcal{E}(\hat{\theta}_1, \kappa_\theta) \triangle \hat{\theta}_2) \oplus (\mathcal{E}(\hat{\theta}_3, \kappa_\theta) \triangle \hat{\theta}_4), \kappa_\theta, \kappa_\theta)$$

where

$$f_\Delta(\theta_a, \theta_b) = \begin{cases} 0, & \theta_a \geq 0 \& \theta_b \geq 0 \\ \theta_a N_\theta, & \theta_a \geq 0 \& \theta_b < 0 \\ \theta_b N_\theta, & \theta_a < 0 \& \theta_b \geq 0 \\ N_\theta (N_\theta + \theta_a + \theta_b), & \theta_a < 0 \& \theta_b < 0. \end{cases}$$

**Algorithm 2** Online Encrypted Learning Algorithm for Global Nash equilibrium

**Initialization**
1. Select initial admissible gain pair $(K_1^{(1)}, K_2^{(1)})$, and a constant $c > 0$ small enough. Let time $t = \tilde{n}$, $T$ be the termination time and $l = 1$.
2. Run the system with initial state $\tilde{x}(0)$ to time $(t - 1)$ under the pair of initial gains.

**Procedure**
1. if $t \leq T$ then
2. Run the system to time $t$ under the pair of latest learned gains $(K_1^{(l)}, K_2^{(l)})$, and the state $\xi(t) = (U \otimes I_N)((0, I_N) \otimes I_N)((I_N + \mathcal{M}_N + 1) \otimes I_N)\tilde{x}(t)$
3. while $\max_i \|P_i^{(l)} - P_i^{(l-1)}\| > c$, for $l > 1$ do
4. for $i = 1, \ldots, N$ do
5. Select $t_1 = t$, and $t_2, \ldots, t_n$ from $\{0, \ldots, t - 1\}$ such that matrix $[\hat{Z}_i(t_1), \ldots, \hat{Z}_i(t_n)]$ is invertible.
6. Quantize the matrices $\hat{Z}_i, \hat{Y}_i$ to $\hat{Z}_i, \hat{Y}_i$ with balance factor $v$ by (41), and reset quantized $\hat{Z}_i$ and $\hat{Y}_i$ to positive integer matrices $\tilde{Z}_i$ and $\tilde{Y}_i$ by (42).
7. Encrypt matrix $\tilde{Z}_i$ by Paillier’s scheme $[\tilde{Z}_i]_{kh} = \mathcal{E}([\tilde{Z}]_{kh}, \kappa_p)$
8. Process encrypted parameter computing in cloud servers, and learn parameters as

$$[\tilde{P}]_{k1} = \oplus_{h=1}^n (\mathcal{E}([\mathcal{Z}]_{kh}, \kappa_p)) \triangle \hat{Y}_{i1}$$

where $k = 1, \ldots, n$.

9. Decrypt parameters computed from cloud servers, and inverse-quantize the parameters with

$$\hat{P}_{i1} = \mathcal{D}([\tilde{P}]_{k1}, \kappa_p, \kappa_s) - \sum_{h=1}^n f_\Delta([\tilde{Z}]_{kh}, \tilde{Y}_{i1})$$

10. end for
11. end
12. Learn the matrices with (37)
13. Update the feedback control gains by (38)
14. Set $t = l + 1$
15. end while
16. Set $t = t + 1$
17. end

**Proof**. According to the designed equation (42), the parameters $\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \hat{\theta}_4$ are reset from $\theta_1, \theta_2, \theta_3$ and $\theta_4$. Consider the additive and multiplicative between parameters, there is

$$(\theta_1 \theta_2) \mod N_\theta = (\hat{\theta}_1 \hat{\theta}_2) \mod N_\theta$$

$$= \mathcal{D}(\mathcal{E}(\hat{\theta}_1, \kappa_\theta) \triangle \hat{\theta}_2, \kappa_\theta, \kappa_\theta) \mod N_\theta.$$
which can further yield
\[(\theta_1, \theta_2) = \mathcal{D}(\hat{\theta}_1, \hat{\theta}_2, \kappa_p, \kappa_s) - f_\Delta(\theta_1, \theta_2). \quad (48)\]

The additive-multiplicative property (46) is not hard to be derived based on (48). This completes the proof. \(\square\)

**Remark 5** The encrypted computing process takes place in cloud servers, and Lemma 3 provides an arithmetic rule for the mixed operation. The additive-multiplicative property (46) is first proposed and effectively guarantees the encrypted additive-multiplicative computation, where the encryption-decryption is embedded in the data transmissions and operations successfully.

The learning process in Algorithm 2 is encrypted for the information security, and solve the global Nash equilibrium securely. Although the errors caused by quantization can not be avoided, as pointed in Remark 4, the implementation can still be guaranteed precise enough when the parameters are quantized by a large enough \(r_m\). The following simulation results will further illustrate the effectiveness of the designed encrypted learning scheme.

### 4 Simulation results

In this section, the effectiveness of the developed privacy information encrypted learning algorithm is shown by a simulation example. We consider the following node dynamics
\[
\begin{align*}
x_1(t + 1) &= Ax_1(t) \\
x_i(t + 1) &= Ax_i(t) + \gamma(t)Bu_i(t) + Dw_i(t) \quad (49)
\end{align*}
\]

where \(A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, B = \begin{bmatrix} -1 \\ 0.5 \end{bmatrix}, \) and \(D = \begin{bmatrix} 0.1 \\ -0.1 \end{bmatrix}\) for \(i = 2, 3, \ldots, 6.\)

![Fig. 3. The communication graph](image)

The communication graph is shown in Fig. 3, and the Laplacian matrix is obtained as (4), where
\[
\begin{align*}
L &= \begin{bmatrix} l_1, l_2, l_3, l_4, l_5 \end{bmatrix}, \quad \alpha &= \begin{bmatrix} -1, 0, -1, 0, 0 \end{bmatrix}^T, \\
l_1 &= \begin{bmatrix} 3, -1, 0, 0, -1 \end{bmatrix}^T, \quad l_2 &= \begin{bmatrix} -1, 2, -1, 0, 0 \end{bmatrix}^T, \\
l_3 &= \begin{bmatrix} 0, -1, 4, -1, -1 \end{bmatrix}^T, \quad l_4 &= \begin{bmatrix} 0, 0, -1, 2, -1 \end{bmatrix}^T, \\
l_5 &= \begin{bmatrix} -1, 0, -1, -1, 3 \end{bmatrix}^T.
\end{align*}
\]

The nonzero eigenvalues of \(L\) are 0.3389, 1.5736, 2.8806, 4.0000, and 5.2068. Suppose that \(\mu = 0.8\), and we choose matrix \(S = I, \eta = 0.2 > \eta^*\), initial gains \(K_1^{(1)} = [1, 0]^T\) and \(K_2^{(1)} = [0, 1]^T\), initial state \(\tilde{x}(0) = [-1, 1, \text{rands}(1, 10)]^T\) in the system.

For the encrypted learning process, we select balance factor \(v = 0.5\), and using the parameter \(r_m = 10^4\) for quantization. Then we use resetting factor \(N_0\) as in (42) to reset matrices to be positive integer parameter ones for operating Paillier’s encryption. After the online encrypted learning and computation at the cyber layer, the encrypted data are sent back from cloud servers, and decrypted to direct the physical layer implementing the control actions. The consensus process of agents is displayed in Fig. 4, where all the states of followers achieve consensus with the leader’s states as desired.

![Fig. 4. The consensus process of agents](image)

![Fig. 5. The evolution of gain parameters](image)

Based on Algorithm 2, the encrypted learning successfully solves the complex graphical game, and the evolution of control gain parameters is displayed by solid lines in Fig. 5, where the gain parameters in the encrypted learning converge to \(K_1 = [-0.1753, 0.1752]\) and \(K_2 = [1.9748, -0.0535]\). To further illustrate the
effect of solutions with quantization errors as pointed in Remark 4, we denote the gain parameters from unprotected computing in Fig. 5 by dotted line as $K_1^* = [-0.1731, 0.1539]$ and $K_2^* = [2.1768, -0.1044]$, which are solved from Algorithm 1 under same initializations. It shows that the developed encrypted learning algorithm achieves a very close solution, and simultaneously the privacy information remains secure, as required.

Some interesting extensions of the current work include, e.g., how to handle attacks over multiple communication channels in realizing Nash equilibrium, how to overcome the errors caused by quantization in encrypted control systems, and how to distinguish particular information from lots of encrypted data in a game.

5 Conclusion

This paper studied the global Nash equilibrium solution of leader-follower multi-agent dynamics associated with a privacy information encrypted learning algorithm. A necessary and sufficient condition for the existence of a global Nash equilibrium is discussed, and the solution of a soft-constrained game is decentralized into the sum of a set of local performance indices. The convergence is guaranteed with an iteratively updated pair of decoupled gains. Based on the proposed additive-multiplicative property and Paillier’s encryption technique, a privacy information encrypted learning algorithm has been designed, upon which the encryption-decryption is embedded in the data transmissions and operations.

Some interesting extensions of the current work include, e.g., how to handle attacks over multiple communication channels in realizing Nash equilibrium, how to overcome the errors caused by quantization in encrypted control systems, and how to distinguish particular information from lots of encrypted data in a game.

Acknowledgement

The authors would like to thank Asso. Prof. Yuan-Hua Ni and Dr. Jianwei Tan for the discussions on this research topic.

Appendix

Proof of Theorem 3. Define $P_i^{(l)} = P_i^{(K_1^{(l)}, K_2^{(l)})}$, $P_i^* = P_i^{(K_1^*, K_2^*)}$, $P_i^{(K_1, K_2)} = P_i^{(K_1, K_2)} = P_i^{(K_1^*, K_2^*)}$, and $P_i^{(K_1, K_2)} = P_i^{(K_1, K_2)} - P_i^{(K_1^*, K_2^*)}$ for all $i = 1, 2, \ldots, N$. Based on Algorithm 1, the solved matrix $P_i^{(l)}$ from the iteration process can generate a set of sequences for $i = 1, \ldots, N$. Let $z \in \mathbb{R}^n$ be any bounded nonzero vector, then the global performance error with any $K_2$ can be written as

\[
J_o(z, K_1^{(l)}, K_2) - J_o(z, K_1^*, K_2^*) = \sum_{i=1}^{N} \left[ z^T \left( P_i^{(K_1^{(l)}, K_2)} - P_i^{(K_1^*, K_2^*)} \right) z \right] \quad (50)
\]

which will be minimized by updating the gain with

\[
K_1^{(l+1)} = \arg \min_{K_1} \left\{ \sum_{i=1}^{N} (z^T P_i^{(K_1, K_2)}) \right\}. \quad (51)
\]

When the global Nash equilibrium exists, based on the Theorem 2, the updated control gain can be obtained by computing the derivation w.r.t. control input as

\[
K_1^{(l+1)} = -\left( \sum_{i=1}^{N} [F_i^{(K_1^{(l)}, K_2)}]^{-1} \sum_{i=1}^{N} [\lambda_i B^T \right.

\times P_i^{(K_1^{(l)}, K_2)} (A - \lambda_i D K_2)] \quad (52)
\]

where $F_i^{(K_1^{(l)}, K_2)} = B^T S B + \lambda_i^2 B^T P_i^{(K_1^{(l)}, K_2)} B$.

Inserting $K_1^{(l+1)}$ into (50) becomes

\[
\sum_{i=1}^{N} (z^T [P_i^{(K_1^{(l+1)}, K_2)}] z) \leq \sum_{i=1}^{N} (z^T [P_i^{(K_1^{(l)}, K_2)}] z) \quad (53)
\]

for any given $K_2$. 
For the second term $\tilde{P}_i^{(K_1, K_2^{(l)})}$ with any $K_1$, there is
\begin{equation}
J_0(z, K_1, K_2^{(l)}) = J_0(z, K_1, K_2^*)
= \sum_{i=1}^{N} \left[ z^T \left( P_i^{(K_1, K_2^{(l)})} - P_i^{(K_1, K_2^*)} \right) z \right]
= \sum_{i=1}^{N} \left[ z^T \left( \tilde{P}_i^{(K_1, K_2^{(l)})} z \right) \right]
\end{equation}
(54)
which can be minimized by
\begin{align*}
K_2^{(l+1)} &= \arg \min_{K_2} \left\{ \sum_{i=1}^{N} [z^T P_i^{(K_1, K_2^{(l)})} z] - \max_{K_2} \sum_{i=1}^{N} [z^T P_i^{(K_1, K_2^*)} z] \right\} \\
&= \arg \max_{K_2} \left\{ \sum_{i=1}^{N} [z^T P_i^{(K_1, K_2^*)} z] \right\}.
\end{align*}
(55)

From computing the derivation w.r.t. disturbance input, we have
\begin{equation}
K_2^{(l+1)} = -\left( \sum_{i=1}^{N} [H_i^{(K_1, K_2^{(l)})}] \right)^{-1} \sum_{i=1}^{N} [\lambda_i D^T P_i^{(K_1, K_2^{(l)})} A + \mu_i B K_1] \right)
\end{equation}
(56)
where $H_i^{(K_1, K_2^{(l)})} = \eta_i^2 D^T D - \lambda_i^2 D^T P_i^{(K_1, K_2^{(l)})} D$.

Inserting $K_2^{(l+1)}$ into (54) yields
\begin{equation}
\sum_{i=1}^{N} (z^T [\tilde{P}_i^{(K_1, K_2^{(l+1)})}] z) \leq \sum_{i=1}^{N} (z^T [\tilde{P}_i^{(K_1, K_2^{(l)})}] z)
\end{equation}
(57)
for any given $K_1$.

According to Lemma 2, (52) and (56) are equal to the updating law (38). Then from the definitions of $\tilde{P}_i^{(K_1, K_2)}$ and $\tilde{P}_i^{(K_1, K_2^{(l)})}$, we can obtain the following inequation
\begin{align*}
\sum_{i=1}^{N} (z^T [\tilde{P}_i^{(K_1^{(l+1)}, K_2^{(l+1)})}] z) \\
= \sum_{i=1}^{N} (z^T [\tilde{P}_i^{(K_1^{(l+1)}, K_2^{(l)})}] + P_i^{(K_1^{(l+1)}, K_2^*)} - P_i^{(K_1^{(l+1)}, K_2^{(l)})}] z) \\
\geq \sum_{i=1}^{N} (z^T [\tilde{P}_i^{(K_1^{(l+1)}, K_2^{(l)})}] z) \\
\end{align*}
(58)
with $\sum_{i=1}^{N} (z^T [P_i^{(K_1^{(l+1)}, K_2^{(l+1)})} - P_i^{(K_1^{(l+1)}, K_2^{(l)})}] z) \geq 0$, and equation
\begin{align*}
\tilde{P}_i^{(K_1^{(l+1)}, K_2^{(l+1)})} \\
= P_i^{(K_1^{(l+1)}, K_2^{(l+1)})} - P_i^{(K_1^{(l+1)}, K_2^{(l)})} \\
= \tilde{P}_i^{(K_1^{(l+1)}, K_2^{(l+1)})} + P_i^{(K_1^{(l+1)}, K_2^*)} - P_i^{(K_1^{(l+1)}, K_2^{(l)})} \\
= \tilde{P}_i^{(K_1^{(l+1)}, K_2^{(l+1)})} + P_i^{(K_1^{(l+1)}, K_2^*)} - P_i^{(K_1^{(l+1)}, K_2^{(l)})} \\
= \tilde{P}_i^{(K_1^{(l+1)}, K_2^{(l+1)})},
\end{align*}
(59)
Let a sequence $\{z^{T\mathbb{P}(1)}z, z^{T\mathbb{P}(2)}z, \ldots, z^{T\mathbb{P}(l)}z, \ldots\}$ be defined with element $z^{T\mathbb{P}(l)}z = \sum_{i=1}^{N} [z^T (P_i^{(l)} - P_i^*) z]$, then it can be written as
\begin{equation}
z^{T\mathbb{P}(l)}z = \sum_{i=1}^{N} z^T [P_i^{(K_1^{(l)}, K_2^*)} - P_i^*] z
\end{equation}
(60)
Inserting $K_1^{(l+1)}$ and $K_2^{(l+1)}$ into equation (60), and using (58) and (59), we obtain
\begin{align*}
z^{T\mathbb{P}(l+1)}z
= \sum_{i=1}^{N} z^T [\tilde{P}_i^{(K_1^{(l+1)}, K_2^{(l+1)})} + \tilde{P}_i^{(K_1^{(l+1)}, K_2^{(l)})}] z \\
\leq \sum_{i=1}^{N} z^T [\tilde{P}_i^{(K_1^{(l+1)}, K_2^{(l)})} + \tilde{P}_i^{(K_1^{(l+1)}, K_2^{(l)})}] z \\
\leq \sum_{i=1}^{N} z^T [\tilde{P}_i^{(K_1^{(l+1)}, K_2^{(l)})} + \tilde{P}_i^{(K_1^{(l+1)}, K_2^{(l)})}] z \\
\leq \sum_{i=1}^{N} z^T [\tilde{P}_i^{(K_1^{(l+1)}, K_2^{(l)})} + \tilde{P}_i^{(K_1^{(l+1)}, K_2^{(l)})}] z \\
= z^{T\mathbb{P}(l)}z.
\end{align*}
(61)
With the initial admissible gains, based on Dini’s theorem (Bartle and Sherbert, 2000), the monotonic non-increasing sequence $\{z^{T\mathbb{P}(l)}z, l = 1, 2, \ldots\}$ will uniform pointwise converge to 0 along with $l \to +\infty$. Based on the definition, there is $\sum_{i=1}^{N} P_i^{(l)} \to \sum_{i=1}^{N} P_i^*$ along with $l \to +\infty$.

According to Lemma 2, the uniform convergence of the sequences $\{K_1^{(l)}, l = 1, 2, \ldots\}$ and $\{K_2^{(l)}, l = 1, 2, \ldots\}$ can be simultaneously achieved with the convergence of sequence $\{P_i^{(l)}, l = 1, 2, \ldots\}$ along with $l \to +\infty$, such that $K_1^{(l)} \to K_1^*$ and $K_2^{(l)} \to K_2^*$. The proof is thus completed. \qed
References

Abu-Khalaf, M., Lewis, F. L., Huang, J., 2006. Policy iterations on the Hamilton–Jacobi–Isaacs equation for $H_\infty$ state feedback control with input saturation. IEEE Transactions on Automatic Control 51 (12), 1989–1995.

Alexandru, A. B., Morari, M., Pappas, G. J., 2018. Cloud-based MPC with encrypted data. In: 2018 IEEE Conference on Decision and Control (CDC). IEEE, pp. 5014–5019.

Aliyu, M., 2011. Nonlinear $H_\infty$-control, Hamiltonian systems and Hamilton-Jacobi equations. CRC Press.

Anderson, B. D., Moore, J. B., 1971. Linear Optimal Control. Prentice-hall.

Bartle, R. G., Sherbert, D. R., 2000. Introduction to real analysis. Vol. 2. Wiley New York.

Başar, T., Bernhard, P., 2008. H-infinity optimal control and related minimax design problems: A dynamic game approach. Springer Science & Business Media.

Bartle, R. G., Sherbert, D. R., 2000. Introduction to real analysis. Vol. 2. Wiley New York.

Başar, T., Bernhard, P., 2008. H-infinity optimal control and related minimax design problems: A dynamic game approach. Springer Science & Business Media.

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

Gu, G., Marinovici, L., Lewis, F. L., 2011. Consensusability of discrete-time dynamic multiagent systems. IEEE Transactions on Automatic Control 57 (8), 2085–2089.

Hadjicostis, C. N., Domínguez-García, A. D., 2020. Privacy-preserving distributed averaging via homomorphically encrypted ratio consensus. IEEE Transactions on Automatic Control 65 (9), 3887–3894.

Han, S., Topcu, U., Pappas, G. J., 2016. Differentially private distributed constrained optimization. IEEE Transactions on Automatic Control 62 (1), 50–64.

Imer, O. C., Yüksel, S., Başar, T., 2006. Optimal control of Itô systems over unreliable communication links. Automatica 42 (9), 1429–1439.

Jiang, J., Jiang, Y., 2020. Leader-following consensus of linear time-varying multi-agent systems under fixed and switching topologies. Automatica 113, 108804.

Jiao, Q., Modares, H., Xu, S., Lewis, F. L., Vannvoudakis, K. G., 2016. Multi-agent zero-sum differential graphical games for disturbance rejection in distributed control. Automatica 69, 24–34.

Kim, J., Lee, C., Shim, H., Cheon, J. H., Kim, A., Kim, M., Song, Y., 2016. Encrypting controller using fully homomorphic encryption for security of cyber-physical systems. IFAC-PapersOnLine 49 (22), 175–180.

Kogiso, K., Fujita, T., 2015. Cyber-security enhancement of networked control systems using homomorphic encryption. In: 2015 54th IEEE Conference on Decision and Control (CDC). IEEE, pp. 6836–6843.

Levis, A. H., Schluter, R. A., Athans, M., 1971. On the behaviour of optimal linear sampled-data regulators. International Journal of Control 13 (2), 343–361.

Liang, W., Chen, H., Zhang, J., Zhao, D., Li, C., 2020. An effective scheme for top-k frequent itemset mining under differential privacy conditions. Sci. China Inf. Sci 63 (5), 1–3.

Lin, H., Su, H., Shi, P., Shu, Z., Lu, R., Wu, Z.-G., 2017. Optimal estimation and control for lossy network: stability, convergence, and performance. IEEE Transactions on Automatic Control 62 (9), 4564–4579.

Lin, H., Su, H., Shu, Z., Wu, Z.-G., Xu, Y., 2015. Optimal estimation in UDP-like networked control systems with intermittent inputs: Stability analysis and suboptimal filter design. IEEE Transactions on Automatic Control 61 (7), 1794–1809.

Liu, J., Huang, J., 2021. Discrete-time leader-following consensus over switching digraphs with general system modes. IEEE Transactions on Automatic Control 66 (3), 1238–1245.

Liu, T., Jiang, Z.-P., 2014. Distributed nonlinear control of mobile autonomous multi-agents. Automatica 50 (4), 1075–1086.

Lopez, V. G., Lewis, F. L., Wan, Y., Liu, M., Hewer, G., Estabridis, K., 2020. Stability and robustness analysis of minmax solutions for differential graphical games. Automatica 121, 109177.

Lu, Y., Zhu, M., 2018. Privacy preserving distributed optimization using homomorphic encryption. Automatica 96, 314–325.

Ma, C.-Q., Zhang, J.-F., 2010. Necessary and sufficient conditions for consensusability of linear multi-agent systems. IEEE Transactions on Automatic Control 55 (5), 1263–1268.

Mo, Y., Murray, R. M., 2016. Privacy preserving average consensus. IEEE Transactions on Automatic Control 62 (2), 753–765.

Moon, J., Barar, T., 2014. Control over lossy networks: A dynamic game approach. In: 2014 American Control Conference. IEEE, pp. 5367–5372.

Murguia, C., Farokhi, F., Shames, I., 2020. Secure and private implementation of dynamic controllers using semihomomorphic encryption. IEEE Transactions on Automatic Control 65 (9), 3950–3957.

Ruan, M., Gao, H., Wang, Y., 2019. Secure and privacy-preserving consensus. IEEE Transactions on Automatic Control 64 (10), 4035–4049.

Sinopoli, B., Spenato, L., Franceschetti, M., Poolla, K., Jordan, M. I., Sastry, S. S., 2004. Kalman filtering with intermittent observations. IEEE transactions on Automatic Control 49 (9), 1453–1464.

Trentelman, H. L., Takaba, K., Monshizadeh, N., 2013. Robust synchronization of uncertain linear multi-agent systems. IEEE Transactions on Automatic Control 58 (6), 1511–1523.

Van Den Broek, W., Engwerda, J., Schumacher, J. M., 2003. Robust equilibria in indefinite linear-quadratic differential games. Journal of Optimization Theory and Applications 119 (3), 565–595.

Wang, Y., Huang, Z., Mitra, S., Dullerud, G. E., 2017. Optimal estimation and control for lossy network: stability, convergence, and performance. IEEE Transactions on Automatic Control 62 (9), 4564–4579.
Xu, L., Mo, Y., Xie, L., 2019. Distributed consensus over markovian packet loss channels. IEEE Transactions on Automatic Control 65 (1), 279–286.
Xu, L., Xiao, N., Xie, L., 2016. Consensusability of discrete-time linear multi-agent systems over analog fading networks. Automatica 71, 292–299.
You, K., Xie, L., 2011. Network topology and communication data rate for consensusability of discrete-time multi-agent systems. IEEE Transactions on Automatic Control 56 (10), 2262–2275.
Zhang, H., Li, L., Xu, J., Fu, M., 2015. Linear quadratic regulation and stabilization of discrete-time systems with delay and multiplicative noise. IEEE Transactions on Automatic Control 60 (10), 2599–2613.
Zhang, H., Wei, Q., Liu, D., 2011. An iterative adaptive dynamic programming method for solving a class of nonlinear zero-sum differential games. Automatica 47 (1), 207–214.
Zhang, R.-R., Guo, L., 2019. Controllability of nash equilibrium in game-based control systems. IEEE Transactions on Automatic Control 64 (10), 4180–4187.
Zhou, K., Doyle, J., Glover, K., 1996. Robust and optimal control. Prentice-hall.
Zhu, Q., Basar, T., 2015. Game-theoretic methods for robustness, security, and resilience of cyberphysical control systems: games-in-games principle for optimal cross-layer resilient control systems. IEEE Control Systems Magazine 35 (1), 46–65.