Nash Equilibrium in a leader-follower multi-agent graphical game against synchronous network attacks

Kun Zhang, Member, IEEE, Ji-Feng Zhang, Fellow, IEEE, Rong Su, Senior Member, IEEE, and Huaguang Zhang, Fellow, IEEE

Abstract—This paper studies the global Nash equilibrium problem of a leader-follower multi-agent graphical game, which yields consensus against organized synchronous network attacks. With the secure hierarchical structure, the consensus condition under active attacks can be analyzed equivalent to the existence and uniqueness of global Nash equilibrium, and the necessary and sufficient condition for the existence of global Nash equilibrium is provided regarding the soft-constrained graphical game. An encrypted learning algorithm is developed to achieve the optimal policies in global Nash equilibrium, and the convergence is guaranteed with an iteratively updated pair of decoupled gains. By using the designed quantization scheme and additive-multiplicative property, encryption-decryption is successfully embedded in the data transmission and computation to overcome the potential passive attacks in unreliable networks. A simulation example is provided to verify the effectiveness of the designed algorithm.

Index Terms—Multi-agent system, Zero-sum game, consensus, network security, iteration learning

I. INTRODUCTION

GAME theory depicts many social scenes in, e.g., economic, military, engineering fields, where 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 the opponent or cooperator can be accessed, there are always uncertainties about the strategies of other partners, no matter how they struggle to offset 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 [1], the soft-constrained game, such as the zero-sum game, can provide an ideal environment to study multi-player decision and control problems, and has successfully developed strategic behavior for some complex dynamics [2]–[4]. However, global Nash equilibrium of a game is very difficult to achieve in the real world, even theoretically, due to the multi-dimensional multi-player networked utility relationships.

With the superiority of handling cooperation/competition behaviors between players in a game, Nash equilibrium in multi-agent systems (MASs) has become a challenging and interesting issue, although MASs have been extensively studied over the past two decades, covering many fundamental cooperative control problems [5]–[8]. For the consensus problem of continuous-time and discrete-time linear MASs, the consensusability and optimal convergence rate, and condition to achieve consensus for multi players/agents, equilibrium. These work successfully settled the foundation of continuous-time and discrete-time linear MASs, the consensusability conditions under undirected communication topology are analyzed and relaxed by [5], [6]. Robust synchronization of MASs is considered in [9], where the additive perturbations are assumed to be either some unknown transfer functions or norm bounded matrices. The effects of undirected graphs on consensusability and optimal convergence rate, and jointly connected switching digraphs are discussed in [10]–[12], where distributed control provides an effective strategy to explore cooperation-competition in a game. Combining the robust consensus and distributed control structure, Nash equilibrium is proposed for MASs in [13], [14], where the minmax strategy must hold for every agent, and it is still a local equilibrium. This work successfully settled the foundation and condition to achieve consensus for multi players/agents, but the global Nash equilibrium in MASs has few research results.

It should be pointed out that, information security is the most crucial condition to achieve Nash equilibrium in a game, where the loss or leakage of one’s situation and strategy may lead to his worst situation. Security control is also an important topic and concern in the human-cyber-physical system, originating in the manufacturing production [15]–[17], where much private information has been exposed to spiteful network users unintentionally, and packet loss from unreliable communication channels has frequented. Along with the development of information technology, 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 potential attackers. Network attacks in recent years have not rarely been reported, such as cyber-attacks on Norsk Hydro
(2019)\textsuperscript{1} and VPNFilter (2018)\textsuperscript{2}, which are two kinds of attacks in active and passive ways. For information or data losses under an active attack, the control and estimation problems are studied in [18]–[20] by modeling the arrival of observations in unreliable communication channels as a random process. A necessary and sufficient consensus condition is discussed in [21] 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 from passive attacks plays a more damaging role in the systems.

To address the situation of unknown leakage of sensitive information, privacy-preserving approaches have recently been proposed for cloud computing [22]–[24], such as differential privacy methods, encrypted computation, and adding artificial noise. Differential privacy presents a quantized performance of privacy preservation, and is successfully utilized in the distributed optimization process [25], which can guarantee user privacy regardless of any auxiliary information that an adversary may have. $\epsilon$-differential privacy is also introduced in [26] to solve the Knapsack problem, where a greedy strategy is proposed. In particular, homomorphic encryption techniques provide an efficient data operation or computation method with partial arithmetic rules \cite{27}–\cite{29}, that are either additive, multiplicative, or even fully homomorphic. A cloud-based model predictive control scheme is discussed with semi-honest servers in [30], where two-party privacy is proposed, and consists of a client-server model and a two-server model. Paillier’s encryption method is developed by [31] 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 systems, which, however, brings many secure technical problems for human-cyber-physical systems in a game.

In this paper, we study the global Nash equilibrium in a leader-follower multi-agent graphical game, which achieves consensus against organized synchronous network attacks. The main contributions can be summarized as follows.

- We have analyzed the consensus problem of leader-follower systems in a soft-constrained game to deal with active attacks, where the consensus is guaranteed by the existence and uniqueness of global Nash equilibrium. To the best of our knowledge, no global Nash equilibrium has been considered from such an overall perspective of MASs.
- This work attempts to solve the global Nash equilibrium by providing an online encrypted learning algorithm, where encryption-decryption is embedded in data transmission and computation to overcome potential passive 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 monotonicity can hardly be obtained in the maxmin problem.

The remainder of this article is organized as follows. First, industrial background and model formulation are presented in Section II and III, respectively. The consensus condition with Nash equilibrium is analyzed, and the encrypted learning algorithm is proposed with convergence and security in Section IV. The simulation results are given in Section V. Finally, conclusions and future work are outlined in Section VI.

## II. Preliminaries

In this paper, the hierarchical structure of human-cyber-physical control systems is considered [16], where the cyber layer directs the physical layer to implement some desired actions or policies and the human layer manages the information process. As presented in Fig. 1, we focus on synchronous network attacks, including active attacks (denial-of-service (DoS), integrity attacks etc.) and passive attacks (sensitive data, privacy attacks etc.), which are generally injected into cloud servers of control systems from the cyber layer and can cause packet loss and privacy violation, respectively.

![The secure hierarchical structure of systems under organized synchronous network attacks](image_url)

In this paper, we study the global Nash equilibrium in a leader-follower multi-agent graphical game, which achieves consensus against organized synchronous network attacks. The main contributions can be summarized as follows.

- We have analyzed the consensus problem of leader-follower systems in a soft-constrained game to deal with active attacks, where the consensus is guaranteed by the existence and uniqueness of global Nash equilibrium. To the best of our knowledge, no global Nash equilibrium has been considered from such an overall perspective of MASs.
- This work attempts to solve the global Nash equilibrium by providing an online encrypted learning algorithm, where encryption-decryption is embedded in data transmission and computation to overcome potential passive 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 monotonicity can hardly be obtained in the maxmin problem.

The remainder of this article is organized as follows. First, industrial background and model formulation are presented in Section II and III, respectively. The consensus condition with Nash equilibrium is analyzed, and the encrypted learning algorithm is proposed with convergence and security in Section IV. The simulation results are given in Section V. Finally, conclusions and future work are outlined in Section VI.

## III. Problem Formulation

Let $\mathcal{V} = \{0, 1, \ldots, N\}$ be the set of $N+1$ agents with $i \in \mathcal{V}$ representing the $i$th agent, and the 0th agent presents the leader. 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

\textsuperscript{1}https://www.hydro.com/en/media/on-the-agenda/cyber-attack/

\textsuperscript{2}https://www.dataprivacyandsecurityinsider.com/2018/05/fbi-warning-russian-hackers-attacking-routers/
agent can receive information from the $j$th agent. The neighborhood of agent $i$ is defined as $N_i = \{j|(j, i) \in \mathcal{E}\}$. The adjacency matrix is defined as $A_{adj} = \{a_{ij}|(N+1) \times (N+1)\}$, 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 $(j, i) \in \mathcal{E}$, else it is directed. The Laplacian matrix $L = [L_{ij}]_{(N+1) \times (N+1)}$ 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 $L = D - A_{adj}$, $D = \text{diag}(d_i)$ is called the in-degree matrix, where $d_i = \sum_{j \in N_i} a_{ij}$.

A directed path on $\mathcal{G}$ from agent $i_1$ to agent $i_k$ is a sequence of ordered edges in the form of $(i_k, i_{k+1}) \in \mathcal{E}$, $k = 1, 2, \ldots, l - 1$. A graph contains a directed spanning tree $T$ if it has at least one agent with directed paths to all other agents.

Suppose that in the multi-agent system, the leader may send information to the followers but does not receive information from any one of the followers, and the communication between the followers is undirected. 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 $x_0(t) \in \mathbb{R}^n$, and assumed to satisfy the dynamic as:

$$x_0(t+1) = Ax_0(t). \quad (1)$$

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

$$x_i(t+1) = Ax_i(t) + B\gamma(t)u_i(t) + Dw_i(t), \quad (2)$$

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 the 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 $i$, respectively, and $i \in \mathcal{V}' = \{1, 2, \ldots, N\}$.

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

$$\mathcal{F}(0) := \{x_i(0), i \in \mathcal{V}'\}, \quad t = 0$$

$$\mathcal{F}(t) := \{\Xi(t), \Gamma(t-1)\}, \quad t = 1, \ldots \quad (3)$$

where $\Xi(t) = (x_1(t), x_2(t), \ldots, x_t(t); i \in \mathcal{V}')$, and $\Gamma(t-1) = (\gamma(0), \gamma(1), \ldots, \gamma(t-1))$. In the leader-follower consensus problem, the Laplacian matrix is

$$L = \begin{bmatrix} 0 & 0^T \\ -\alpha & L \end{bmatrix} \quad (4)$$

where $0_N = [0, 0, \ldots, 0]^T$, $\alpha = [a_{10}, a_{20}, \ldots, a_{NN}]^T$, and

$$L = \begin{bmatrix} \sum_{j \in N_1} a_{1j} & -a_{12} & \cdots & -a_{1N} \\ -a_{21} & \sum_{j \in N_2} a_{2j} & \cdots & -a_{2N} \\ \vdots & \ddots & \ddots & \vdots \\ -a_{N1} & -a_{N2} & \cdots & \sum_{j \in N_N} a_{Nj} \end{bmatrix}.\quad (5)$$

**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 the 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 of packet loss $\{\gamma(t)\}_{t \geq 0}$ in the channel is distributed as an i.i.d. Bernoulli process with probability $\mathbb{P}(\gamma(t) = 1) = \mu$.

**Remark 2.** System (2) is considered constructing potential active attackers, where the packet losses are caused by a malicious jammer and can be considered identical in the system [21], [32]. In addition, the follower’s dynamic (2) with packet loss information (3) can model the TCP-like network protocol with DoS attacks [16], [33].

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 $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. For a soft-constrained dynamic game, the consensus and disturbance policies for followers are considered by

$$u_i(t) = K_{i1} \sum_{j \in N_i} a_{ij}(x_j(t) - x_i(t)),$$

$$w_i(t) = -K_{i2} \sum_{j \in N_i} a_{ij}(x_j(t) - x_i(t)) \quad (5)$$

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 $\bar{x}(t) = [x_0(t)^T, x_1(t)^T, \ldots, x_N(t)^T]^T$, then, from dynamics (1) and (2), the system can be rewritten as

$$\bar{x}(t+1) = (I_{N+1} \otimes A + \gamma(t)L \otimes BK_1 - L \otimes DK_2)\bar{x}(t) \quad (6)$$

where $\otimes$ means Kronecker product.

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

$$\delta(t+1) = [(I_{N+1} - \mathcal{M}_{N+1}) \otimes I_n] \bar{x}(t+1) = (I_{N+1} \otimes A + \gamma(t)L \otimes BK_1 - L \otimes DK_2)\delta(t). \quad (7)$$

Consider $\bar{\delta}(t) = [(0_{N,1}, I_N) \otimes I_n] \delta(t) = [\bar{\delta}_2(t), \ldots, \delta_{N+1}(t)]^T$, one has

$$\bar{\delta}(t+1) = (I_{N} \otimes A + \gamma(t)L \otimes BK_1 - L \otimes DK_2)\bar{\delta}(t), \quad (8)$$

and it follows that $\bar{\delta}(t) = 0$ if and only if $x_0(t) = x_1(t) = \cdots = x_N(t)$.

**Definition 1.** [34] System (8) is said to have $L_2$-gain less than or equal to $\eta$, if for any initial state $\delta(0)$ and $z(t) = (L \otimes DK_2)\delta(t)$, the response $z(t) = C(t)\delta(t)$ corresponding to any $w(t) \in L_2[0, \infty)$ satisfies:

$$\mathbb{E}\left[\sum_{t=0}^\infty ||z(t)||^2 \mathcal{F}(0)\right] \leq \eta^2 \mathbb{E}\left[\sum_{t=0}^\infty ||w(t)||^2 + \phi(\delta(0))\right] \quad (9)$$

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

**Bounded $L_2$-gain consensus problem:** [35] 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...
$w_1(t) = 0$ and satisfy the $L_2$-gain condition (9) for a given $\eta$ when $w_i(t) \neq 0$, with initial state $\delta(0) = 0$.

To achieve condition (9), the soft-constrained dynamic game can be utilized to construct the $H_\infty$ optimal control [36]. We define that symmetrical positive definite matrices, $\kappa$, $\phi$ the least common multiple, $\phi$ 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 $L$. Note that the matrix $Q_N$ is positive definite with the undirected communication between the followers.

Based on information structure (3) and optimal control theory [38], [39], we can rewrite the steady-state performance index (10) in an infinite horizon as $J_o(\delta(t), K_1, K_2) = \delta(t)T\mathcal{P}(t)$, $\mathcal{P} = \mathcal{P}^T > 0$, and further obtain

$$J_o(\delta(t), K_1, K_2) = \mathcal{E}\{\delta(t)^T \mathcal{Q}_{(K_1, K_2)}(t) \delta(t) | \mathcal{F}(t)\}$$

(10)

where $\mathcal{Q}_{(K_1, K_2)}(t) = Q_N \otimes I_N + \gamma(t)(L \otimes BK_1)^T \mathcal{S}(L \otimes BK_1) - \eta^2(L \otimes DK_2)^T \mathcal{S}(L \otimes DK_2)$, $\gamma(t)$ 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 $L$. Note that the matrix $Q_N$ is positive definite with the undirected communication between the followers.

Based on information structure (3) and optimal control theory [38], [39], we can rewrite the steady-state performance index (10) in an infinite horizon as $J_o(\delta(t), K_1, K_2) = \delta(t)T\mathcal{P}(t)$, $\mathcal{P} = \mathcal{P}^T > 0$, and further obtain

$$J_o(\delta(t), K_1, K_2) = \mathcal{E}\{\delta(t)^T \mathcal{Q}_{(K_1, K_2)}(t) \delta(t) | \mathcal{F}(t)\}$$

(11)

with the system dynamic (8).

Then, the global Nash equilibrium solution $(K_1^*, K_2^*)$ of this zero-sum game satisfies

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

$$= \max_{K_2} \min_{K_1} J_o(\delta(0), K_1, K_2)$$

(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 consensus 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

$$J_o(\delta(t), K_1^*, K_2^*) = \mathcal{E}\{\delta(t)^T \mathcal{Q}_{(K_1^*, K_2^*)}(t) \delta(t) | \mathcal{F}(t)\}$$

+ $J_o(\delta(t+1), K_1^*, K_2^*) | \mathcal{F}(t)\}$,

(13)

which gives

$$\max_{K_2} \left\{ \mathcal{E}\{J_o(\delta(t+1), K_1^*, K_2) - J_o(\delta(t), K_1^*, K_2) + \delta(t)T \mathcal{Q}_{(K_1^*, K_2)}(t) \delta(t) | \mathcal{F}(t)\} \right\} = 0.$$  

(14)

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

$$\mathcal{E}\{J_o(\delta(t_s + 1), K_1^*, K_2) - J_o(\delta(0), K_1^*, K_2) | \mathcal{F}(0)\}$$

+ $\mathcal{E}\{\sum_{t=0}^{t_s} \delta(t)^T \mathcal{Q}_{(K_1^*, K_2)}(t) \delta(t) | \mathcal{F}(0)\} \leq 0,$

(15)
Let \( \bar{\delta}(0) = 0 \), we have
\[
E\left(\sum_{t=0}^{\infty} ||z(t)||^2 |\mathcal{F}(0)|\right) \leq \eta^2 \sum_{t=0}^{\infty} ||z(t)||^2 .
\] (16)

(Sufficient) Considering that condition (9) holds, the global performance index (10) is bounded with admissible controls as \( J_o(\bar{\delta}(t), K_1, K_2) \) \( < \infty \). The function \( J_o(\bar{\delta}(t), K_1, K_2) \) is strictly convex with respect to (w.r.t) \( (I \otimes K_1)\bar{\delta}(t) \) and strictly concave w.r.t \( (I \otimes K_2)\bar{\delta}(t) \). Based on Theorem 2.3 by [36], 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 the global Nash equilibrium, the scheme is detailed as 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 \( i \) in \( \mathcal{V}' \) as
\[
J_i(\xi_i(t), K_1, K_2) = E\left(\sum_{k=t}^{\infty} \xi_i(k)^T \bar{Q}(t) \xi_i(k) |\mathcal{F}(t)\right) \quad (17)
\]
for the local dynamic
\[
\dot{\xi}_i(t) = (A + \gamma(t)\lambda_i BK_1 - \lambda_i DK_2) \xi_i(t) \quad (18)
\]
where \( \bar{Q}(t) = I + \gamma(t)(BK_1)^T S(BK_1) - \eta^2(DK_2)^T(DK_2) \).

**Proof.** Consider a pair \((K_1^*, K_2^*)\) being the solution of global Nash equilibrium, which holds for the following optimization problem
\[
\min_{K_1 \otimes K_2} \max_{\bar{\delta}(t), K_1, K_2} J_o(\bar{\delta}(t), K_1, K_2) \quad (19)
\]
s.t. \( \bar{\delta}(t+1) = (I_N \otimes A + \gamma(t)L \otimes BK_1 - L \otimes DK_2)\bar{\delta}(t) \).

Let \( U \in \mathbb{R}^{N \times N} \) be such a unitary matrix that \( U^TU = I_N \), and \( UU^T = \Lambda \equiv \text{diag}(\lambda_1, \ldots, \lambda_N) \). By choosing \( \xi(t) = [\xi_1(t), \ldots, \xi_N(t)] = (U \otimes I_n)\bar{\delta}(t) \), it is not difficult to obtain that
\[
J_o(\bar{\delta}(t), K_1, K_2) = E\left(\sum_{k=t}^{\infty} \xi_i(k)^T(U \otimes I_n)\left[Q_N \otimes I_n + \gamma(t)
\right. \times (L \otimes BK_1)^T S(L \otimes BK_1) - \eta^2(L \otimes DK_2)^T S(L \otimes DK_2)\right]
\times (U \otimes I_n)\xi(k) |\mathcal{F}(t)\}
\]
\[
= E\left(\sum_{k=t}^{\infty} \sum_{i=1}^{N} \xi_i(k)^T \lambda_i^2 I_n + \gamma(t)\lambda_i^2(BK_1)^T S(BK_1)
\times (BK_1) - \eta^2\lambda_i^2(DK_2)^T(DK_2)\right) |\mathcal{F}(t)\}
\]
\[
= \sum_{i=1}^{N} \lambda_i^2 J_i(\xi_i(t), K_1, K_2).
\] (20)

For the dynamic (8), it becomes
\[
\dot{\xi}(t+1) = \left( I_N \otimes A + \gamma(t)\Lambda \otimes BK_1 - \Lambda \otimes DK_2 \right) \xi(t),
\] (21)
which is equivalent to the simultaneous dynamic
\[
\xi_i(t+1) = (A + \gamma(t)\lambda_i BK_1 - \lambda_i DK_2) \xi(t) \quad (22)
\]
for all \( i \in \mathcal{V}' \).

Based on the transform by \( \bar{\delta}(t) = (U \otimes I_n)^T \xi(t) \), problem (19) is equivalent to the optimization
\[
f_1 : \min_{K_1} \max_{K_2} \sum_{i \in \mathcal{V}'} \lambda_i^2 J_i(\xi_i(t), K_1, K_2) \quad (23)
\]
s.t. \( \xi_i(t+1) = (A + \gamma(t)\lambda_i BK_1 - \lambda_i DK_2) \xi_i(t), \forall i \in \mathcal{V}' \).

This completes the proof. \( \square \)

From 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 [39], 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 \( J_i(\xi_i(t), K_1, K_2) = \xi_i^T(t)P_i\xi_i(t) \). By inserting (18) into (17), we have
\[
J_i(\xi_i(t), K_1, K_2) = \xi_i^T(t)P_i\xi_i(t)
\]
\[
= E\{\xi_i(k)^T \bar{Q}(t) \xi_i(k) |\mathcal{F}(t)\}
\]
\[
+ E\{J_i(\xi_i(t+1), K_1, K_2) |\mathcal{F}(t)\}
\]
\[
= \xi_i^T(t)[I - \eta^2(DK_2)^T(DK_2) + (A + \mu_1 BK_1 - \lambda_i DK_2)^T(P_i + \mu_1 BK_1 - \lambda_i DK_2)\]
\[
+ \mu_1(BK_1)^T[S + (1 - \mu_2^2)P_i](BK_1)]\xi_i(t).
\] (24)

**Theorem 2.** The consensus condition (9) of multi-agent system (8) can be guaranteed if and only if (a) the condition \( \sum_{i=1}^{N} H_i > 0 \); or (b) the pair of optimal gains satisfies
\[
\begin{cases}
K_1^* = -(\sum_{i=1}^{N} F_i)^{-1} \sum_{i=1}^{N} \lambda_i B^TP_i(A - \lambda_i DK_2^*) \\
K_2^* = -(\sum_{i=1}^{N} H_i)^{-1} \sum_{i=1}^{N} \lambda_i D^TP_i(A + \mu_1 BK_1^*)
\end{cases}
\] (25)
where \( F_i = B^T S + \lambda_i^2 B^TP_iB \) and \( H_i = \eta^2 D^TD - \lambda_i^2 D^TP_iD \).

**Proof.** (Necessary) Define \((K_1^*, K_2^*)\) as the solution of global Nash equilibrium, using Lemma 1, we solve optimization (23) by summing the derivative of \( J_i(\xi_i(t), K_1, K_2) \) w.r.t control input \( u_i(t) = K_{i1}\xi_i(t) \) and disturbance input \( w_i(t) = K_{i2}\xi_i(t) \), respectively, and it can be easily obtained the optimal gains satisfying
\[
\sum_{i=1}^{N} F_i K_{i1}^* = -\sum_{i=1}^{N} \lambda_i B^TP_i(A - \lambda_i DK_2^*)
\]
\[
\sum_{i=1}^{N} H_i K_{i2}^* = -\sum_{i=1}^{N} \lambda_i D^TP_i(A + \mu_1 BK_1^*).
\] (26)

Thus, based on Theorem 1, the consensus condition (9) is guaranteed by the existence and uniqueness of global Nash equilibrium (12), which requires \( \sum_{i=1}^{N} H_i > 0 \) and gives (25).
Remark 3. From Lemma 1, the global Nash equilibrium is a necessary but possibly insufficient condition for the simultaneous local Nash equilibriums with the following optimization

\[ \min_{K_1, K_2} \max_{i} J_i(\xi_i(t), K_1, K_2) \]

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

for each i. It can also be easily obtained as \( H_i > 0 \Rightarrow (\sum_{i=1}^{N} H_i) > 0 \) and \( H_i > 0 \Leftrightarrow (\sum_{i=1}^{N} H_i) > 0 \). Different with the typical consensus in a distributed zero-sum game [13], [14], we do not require that the local Nash equilibriums hold simultaneously.

In this subsection, the consensus condition of multi-agent system is proven to be equivalent to the existence and uniqueness of the global Nash equilibrium (12). The necessary and sufficient condition of consensus is discussed in Theorem 2, and the decentralized dynamic in Lemma 1 actually provides a medium for solving the global Nash equilibrium and achieving the consensus.

B. The solution of global Nash equilibrium

To further solve the global Nash equilibrium from the set of decentralized local performance indices, we decouple the pair of gains (25) in the following lemma.

Lemma 2. The coupled control gains with the form (25) for the global Nash equilibrium can be decoupled as

\[
K_1^* = -M_1^{-1}(\sum_{i=1}^{N} F_i)^{-1}[\sum_{i=1}^{N} \lambda_i B^T P_i A]
+ \sum_{i=1}^{N}[\lambda_i^2 B^T P_i D][\sum_{i=1}^{N} H_i]^{-1}[\sum_{i=1}^{N} \lambda_i D^T P_i A];
\]

\[
K_2^* = -M_2^{-1}(\sum_{i=1}^{N} H_i)^{-1}[\sum_{i=1}^{N} \lambda_i D^T P_i A]
- \sum_{i=1}^{N}[\mu \lambda_i D^T P_i B][ \sum_{i=1}^{N} F_i]^{-1}[\sum_{i=1}^{N} \lambda_i B^T P_i A]
\]

where \( M_1 = I_{m_1} + (\sum_{i=1}^{N} F_i)^{-1}[\sum_{i=1}^{N} \lambda_i B^T P_i D][\sum_{i=1}^{N} H_i]^{-1}
\]

\[ \sum_{i=1}^{N}[\lambda_i^2 D^T P_i B], \]

\[ \sum_{i=1}^{N} F_i [\sum_{i=1}^{N} \lambda_i^2 B^T P_i D], \]

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 (29) based on (25), when the global Nash equilibrium exists. It gives the proof.

Combining Theorem 2, the decoupled optimal gain pair (29) satisfies the set of decentralized local performance indices (24) 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

\[
f_2(P_i^{(l)}) = I_{m_1} + [\mathcal{Y}^{(l)}]^{-1} \sum_{i=1}^{N} \lambda_i^2 B^T P_i^{(l)} D][\mathcal{H}^{(l)}]^{-1}
\times \sum_{i=1}^{N}[\mu \lambda_i^2 D^T P_i^{(l)} B] ;
\]

\[
f_3(P_i^{(l)}) = I_{m_2} + [\mathcal{H}^{(l)}]^{-1} \sum_{i=1}^{N}[\lambda_i^2 D^T P_i^{(l)} B][\mathcal{Y}^{(l)}]^{-1}
\times \sum_{i=1}^{N}[\lambda_i^2 B^T P_i^{(l)} D];
\]

\[
f_4(P_i^{(l)}) = -M_1^{-1}[\mathcal{Y}^{(l)}]^{-1}[\sum_{i=1}^{N} \lambda_i B^T P_i^{(l)} A]
+ \sum_{i=1}^{N}[\lambda_i^2 B^T P_i^{(l)} D][\mathcal{H}^{(l)}]^{-1}[\sum_{i=1}^{N} \lambda_i D^T P_i^{(l)} A];
\]

\[
f_5(P_i^{(l)}) = -M_2^{-1}[\mathcal{H}^{(l)}]^{-1}[\sum_{i=1}^{N} \lambda_i D^T P_i^{(l)} A]
- \sum_{i=1}^{N}[\mu \lambda_i D^T P_i^{(l)} B][\mathcal{Y}^{(l)}]^{-1}[\sum_{i=1}^{N} \lambda_i B^T P_i^{(l)} A]
\]

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

Algorithm 1 Unprotected Decentralized Iterative Learning for Global Nash equilibrium

Initialization

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

Procedure

1: \( \text{while} \ (\max_i \| P_i^{(l)} - P_i^{(l-1)} \| > \epsilon) \& (l > 1) \) \( \text{do} \)
2: \( \text{for} \ i = 1, \ldots, N \ \text{do} \)
3: \( \triangleright \text{Solve local matrix } P_i^{(l)} \text{ from equation (24) with the gain pair } (K_1^{(l)}, K_2^{(l)}) \)
4: \( \triangleright \text{Tune the parameters as} \)
5: \( F_i^{(l)} = [B^T S (B) + \lambda_i^2 B^T P_i^{(l)} B]; \)
6: \( H_i^{(l)} = [\eta^2 D^T D - \lambda_i^2 D^T P_i^{(l)} D] \)

3: \( \text{end} \)
4: \( \text{end} \)
5: \( \text{end} \)
6: \( \circ \text{Learn the matrices with} \)
7: \( M_1^{(l)} \leftarrow f_2(P_i^{(l)}); M_2^{(l)} \leftarrow f_3(P_i^{(l)}) \)
8: \( \circ \text{Update the control gains by} \)
9: \( K_1^{(l+1)} \leftarrow f_4(P_i^{(l)}); K_2^{(l+1)} \leftarrow f_5(P_i^{(l)}) \)
10: \( \circ \text{Set} \ l = l + 1 \)
11: \( \text{end while} \)

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

Proof. Define \( P_i^{(l)} = P_i^{(K_1^{(l)}, K_2^{(l)})}, P_i^{*} = P_i^{(K_1^*, K_2^*)} \), \( \tilde{P}_i^{(K_1, K_2)} = P_i^{(K_1, K_2)} - P_i^{(K_1^*, K_2^*)} \), and \( \tilde{P}_i^{(K_1^*, K_2^*)} \) with the number of iterations \( l = 1,2, \ldots \), then there is

\[
\sum_{i=1}^{N} (P_i^{(l)} - P_i^{*}) = \sum_{i=1}^{N} (P_i^{(K_1^{(l)}, K_2^{(l)})} + \tilde{P}_i^{(K_1^*, K_2^*)}).
\]

(33)

[Monotonicity]. The main idea of monotonicity proof is as follows, and the detail is given in Appendix.

In the iterations, for any bounded nonzero vector \( z \in \mathbb{R}^n \), the sequence \( \{ z^T \tilde{P}_i^{(l)} z \} \), \( \tilde{P}_i^{(l)} = \sum_{i=1}^{N} [P_i^{(l)} - P_i^{*}] \), is monotonic nonincreasing from the following steps.

Step 1: The equation (32) updates \( K_1^{(l)} \) to \( K_1^{(l+1)} \), which minimize the term \( \sum_{i=1}^{N} z^T [P_i^{(l)} - P_i^{*}] z \).

Step 2: Gain \( K_2^{(l)} \) is also updated to \( K_2^{(l+1)} \), which minimize \( \sum_{i=1}^{N} z^T [P_i^{(K_1^{(l+1)}, K_2^{(l)})} + \tilde{P}_i^{(K_1^*, K_2^*)}] z \).

Step 3: We compare these terms, and can obtain

\[
z^T \tilde{P}_i^{(l)} z \geq \sum_{i=1}^{N} [z^T (P_i^{(l)} - P_i^{*}) z] \geq \sum_{i=1}^{N} [z^T (P_i^{(l)} - P_i^{*}) z]
\]

(34)

for all \( l \).

[Boundedness]. The inequations hold as:

\[
z^T \tilde{P}_i^{(l)} z \geq z^T \tilde{P}_i^{(l)} z \geq \cdots \geq z^T \tilde{P}_i^{(l)} z \geq 0
\]

(35)

within the interval \([0, z^T \tilde{P}_i^{(l)} z] \).

Combining the monotonicity and boundedness, based on Dini’s theorem [40], the monotonic nonincreasing sequence \( \{ z^T \tilde{P}_i^{(l)} z \} \) will uniform pointwise converge to 0 along with \( l \to +\infty \), which indicates \( \sum_{i=1}^{N} P_i^{(l)} \to \sum_{i=1}^{N} P_i^{*} \) as \( l \to +\infty \). According to Lemma 2, the uniform convergence of the sequences \( \{ K_1^{(l)} \} \) and \( \{ K_2^{(l)} \} \) can be simultaneous achieved with the convergence of sequence \( \{ P_i^{(l)} \} \) 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.

In Algorithm 1, the set of decentralized local performance indices (24) and the decoupled pair of gains (29) provide an effective way to solve the global Nash equilibrium for the leader-follower system. Theorem 3 successfully gives convergence with an iterative learning process, where due to the competitive relationship of policies in the soft-constrained zero-sum game, the convergence of the maxmin solution is generally hard to obtain.

C. The online encrypted learning algorithm

The procedure in Algorithm 1 involves many unprotected private data in the networked system. To further guarantee the privacy security of states and policies under potential passive attackers, the data transmissions and operations should be encrypted between the cyber layer and physical layer, including iterative computations of decentralized equation (27) in the cloud servers. The implementation of a developed encrypted learning process is detailed in the following.

Based on equation (24), we have

\[
\tilde{z}_i(t)^T P_i(t) \xi_i(t) = \tilde{z}_i(t)^T [I + \mu(B K_1)^T S(B K_1) - \eta^2 (D K_2)^T] \times (D K_2) \xi_i(t) + \xi_i(t + 1) P_i(t + 1) \xi_i(t + 1)
\]

(36)

which can be vectored as

\[
\tilde{z}_i(t)^T [I + \mu(B K_1)^T S(B K_1) - \eta^2 (D K_2)^T (D K_2)] \xi_i(t)
\]

(37)

\[
\tilde{z}_i(t)^T [I + \mu(B K_1)^T S(B K_1) - \eta^2 (D K_2)^T (D K_2)] \xi_i(t)
\]

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

Note that in equation (37), \( \tilde{z}_i(t) \) is a vector consisting of consensus states directly, which are sensitive data in privacy preservation; vector \( \tilde{P}_i \) is unknown, needs to be solved, and is also privacy as it provides the core control parameters for the consensus process; \( Y_i(t) \) is a measured utility value in practice. With enough sampling data, equation (37) can be further augmented as \( \tilde{P}_i = Z_i^T Y_i \), 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. Denoting \([\cdot]_{kh} \) as the matrix element of row \( k \) and column \( h \), the online encrypted learning process is developed in Algorithm 2.

Lemma 3. Let \( \tilde{P}_i = g_1(r_{m,v}) (Z_i) \cdot g_2(r_{m,v}) (Y_i) \), where mappings \( g_1(r_{m,v})(\cdot) : \mathbb{R}^{n \times \bar{n}} \to \mathbb{R}^{n \times \bar{n}} \) and \( g_2(r_{m,v})(\cdot) : \mathbb{R}^{n \times \bar{n}} \to \mathbb{Z}^{\bar{n}} \) are with a factor pair \( (r_{m,v}) \). For any \( \varepsilon > 0 \), we can always find \( g_1(r_{m,v})(\cdot) \) and \( g_2(r_{m,v})(\cdot) \), such that \( ||\tilde{P}_i - \bar{P}_i|| \leq \varepsilon \).

Proof. Consider a parameter \( \rho \in \mathbb{R} \), there are always \( r_1, r_2 \in \mathbb{N} \), such that \( \rho \leq 0 \leq \frac{r_1 + r_2}{2} \), which is proved in [29]. For the elements in matrices \( Z_i \) and \( Y_i \), a common multiple \( m \in \mathbb{N} \) can be found to quantize these elements by

\[
q_1(\rho) = \lceil r_m \rho \rceil
\]

(41)

with a max quantization error \( 1/r_m \). Define \( g_1(r_{m,v})(Z_i) = \frac{1}{m} r_m Z_i \) and \( g_2(r_{m,v})(Y_i) = \frac{1}{v r_m} [r_m Y_i] \), we have

\[
||\tilde{P}_i - \bar{P}_i|| \leq \frac{\bar{n} \sqrt{n}}{r_m}
\]

(42)

for any fixed \( \bar{n} \), then, we select \( r_m \geq \sqrt{\frac{\bar{n} \sqrt{n}}{\varepsilon}} \), there is always \( ||\tilde{P}_i - \bar{P}_i|| \leq \varepsilon \). The proof is completed.

Remark 4. As pointed out in [27], the error caused by quantization is still open in changing the performance of the networked control system. Different from the common fixed-point arithmetic [31], [37], the factor \( v \) here can mitigate the influence of quantization errors, overflow or underflow in computing processes, moreover, the influence will be not accumulated during the convergence.
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 small enough constant \(\varepsilon > 0\). Set initial time \(t = \bar{n}\), termination time \(\mathcal{T}\) and number of iterations \(l = 1\);
2: Run the system from time 0 to \((t - 1)\) with initial state \(\bar{x}(0)\) and gain pair

Procedure
1: if \(t \leq \mathcal{T}\) then
2: Run the system to time \(t\) with the latest learned gain pair \((K_1^{(l)}, K_2^{(l)})\), and obtain a new state \(\xi(t) = (U \otimes I_n)([0_N, I_N] \otimes I_n)([I_{N+1} - (\Xi N + 1) \otimes I_n])\bar{x}(t)\)
3: while \((\max_i \| P_i^{(l)} - P_i^{(l-1)} \| > \varepsilon) \& (l > 1)\) do
4: for \(i = 1, \ldots, N\) do
5: \(\triangleright \) Select \(t_i = t, t_2, \ldots, t_i\) from \(\{0, \ldots, t-1\}\) such that matrix \([\bar{Z}_i(t_1), \ldots, \bar{Z}_i(t_i)]\) is invertible
6: \(\triangleright \) Quantize and reset matrices \(\bar{Z}_i\) and \(\bar{Y}_i\) by (41) and (43) to positive integer matrices \(\bar{Z}_i\) and \(\bar{Y}_i\)
7: \(\triangleright \) Encrypt matrices \(\bar{Z}_i\) by Paillier’s scheme \([\bar{Z}_i]_{kh} = \mathcal{E}(\bar{Z}_i_{kh}, \kappa_p)\) (38)
8: \(\triangleright \) Process encrypted parameter learning in cloud servers, and compute parameters as \([\bar{P}_i]_{1k} = \oplus_{h=1}^{n} (\mathcal{E}(\bar{Z}_i^T_{kh}, \kappa_p) \Delta \bar{Y}_i_{kh})\) (39)
9: \(\triangleright \) Decrypt and inversely quantize parameters by (44) with \([\bar{P}_i]_{1k} = \mathcal{D}(\bar{P}_i_{1k}, \kappa_p, \kappa_a) - \sum_{h=1}^{n} f\Delta([\bar{Z}_i]_{kh}, [\bar{Y}_i]_{kh})\)
10: \(\triangleright \) Tune the parameters as (30)
end for
11: o Learn the matrices with (31)
12: o Update the feedback control gains by (32)
13: o Set \(l = l + 1\)
end while
16: Set \(t = t + 1\)
end if

End Procedure

In the encrypted learning process, we use Paillier’s encryption method to encrypt the data before transmissions and operations, and the detailed privacy data flow in the encrypted learning architecture is presented in Fig. 2.

According to Lemma 3, the matrices \(\bar{Z}_i\) and \(\bar{Y}_i\) are quantized to integer parameter matrices \(\bar{Z}_i = q_1(\bar{Z}_i)\) and \(\bar{Y}_i = q_1(\bar{Y}_i)\), and for any parameter \(\rho\)’ in matrix \((\bar{Z}_i, \bar{Y}_i)\), it can also be inversely quantized as \(q_2(\rho) = \rho / r_2^m\). To make sure that all the computing parameters are positive integers, we reset parameter \(\bar{\theta} \in \bar{Z}_i \cup \bar{Y}_i\) by

\[
\bar{\theta} = \begin{cases} 
\theta + N_{\theta}, & \theta < 0 \\
\theta, & \theta \geq 0
\end{cases}
\]  
(43)

where \(\bar{Z}_i \cup \bar{Y}_i\) denotes the set of parameters in \(\bar{Z}_i\) and \(\bar{Y}_i\), \(N_{\theta} = 2\theta_m + 1\) is the resetting factor, and \(\theta_m = \max(\theta, \theta \in \text{abs}(\bar{Z}_i \cup \bar{Y}_i))\). Then, matrices \(\bar{Z}_i\) and \(\bar{Y}_i\) can be easily reset to the positive integer matrices \(\bar{Z}_i\) and \(\bar{Y}_i\) by using (43).

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

**Lemma 4.** For any parameters \(\theta_1, \theta_2, \theta_3, \theta_4 \in \bar{Z}_i \cup \bar{Y}_i\), the corresponding positive integer parameters are \(\bar{\theta}_1, \bar{\theta}_2, \bar{\theta}_3, \bar{\theta}_4 \in \bar{Z}_i \cup \bar{Y}_i\) from equation (43), then, to implement Paillier’s semi-homomorphic encryption, we give the following additive-multiplicative property

\[
(\theta_1 \theta_2 + \theta_3 \theta_4) = \mathcal{D}(\mathcal{E}(\bar{\theta}_1, \kappa_p) \Delta \bar{\theta}_2 + \mathcal{E}(\bar{\theta}_3, \kappa_p) \Delta \bar{\theta}_4), \kappa_p, \kappa_a) 
\]  
(44)

where

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

**Proof.** According to the designed equation (43), the parameters \(\bar{\theta}_1, \bar{\theta}_2, \bar{\theta}_3, \bar{\theta}_4\) are reset from \(\theta_1, \theta_2, \theta_3, \theta_4\). Considering the additive and multiplicative between parameters, there is

\[
(\theta_1 \theta_2) \mod N_{\theta} = (\bar{\theta}_1 \bar{\theta}_2) \mod N_{\theta} = \mathcal{D}(\mathcal{E}(\bar{\theta}_1, \kappa_p) \Delta \bar{\theta}_2, \kappa_p, \kappa_a) \mod N_{\theta},
\]

which can further yield

\[
(\theta_1 \theta_2) = \mathcal{D}(\mathcal{E}(\bar{\theta}_1, \kappa_p) \Delta \bar{\theta}_2, \kappa_p, \kappa_a) - f\Delta(\theta_1, \theta_2). 
\]  
(46)

The additive-multiplicative property (44) can be guaranteed based on (46). This completes the proof.

**Remark 5.** The encrypted computing process takes place in cloud servers, and Lemma 4 provides an arithmetic rule for the mixed operation. The additive-multiplicative property (44) 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 following simulation results will further illustrate the effectiveness of the designed encrypted learning scheme.
V. SIMULATION RESULTS

In this section, the consensus of leader-follower MASs against synchronous network attacks is verified by using the developed scheme. We consider the following MASs’ dynamic

\[
\begin{align*}
x_0(t+1) &= Ax_0(t)
x_i(t+1) &= Ax_i(t) + \gamma(t)Bu_i(t) + Dw_i(t)
\end{align*}
\]

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 \( i = 1, 2, \ldots, 5 \). The communication graph is shown in Fig. 3, and the corresponding Laplacian matrix can be obtained as (4), where \( L = \begin{bmatrix} 0 & 0.1 \\ -0.1 & 0 \end{bmatrix} \), \( L = \begin{bmatrix} l_1, l_2, l_3, l_4, l_5 \end{bmatrix} \), \( \alpha = [-1, 0, -1, 0, 0]^T \), \( l_1 = [3, -1, 0, 0, -1]^T \), \( l_2 = [-1, 2, -1, 0, 0]^T \), \( l_3 = [0, -1, 4, -1, -1]^T \), \( l_4 = [0, 0, -1, 2, -1]^T \), \( l_5 = [-1, 0, -1, -1, 3]^T \), and the nonzero eigenvalues of \( L \) can be calculated as 0.3389, 1.5736, 2.8806, 4.0000, and 5.2068.

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

Based on the proposed Algorithm 2, the consensus process of agents is displayed in Fig 4, where all the states of followers achieve consensus against the active attacks, as desired. For the passive attacks, we compare the encrypted data computed from (39) in cloud servers with the decrypted actual parameters from (40), as presented in Table I, where the encrypted data are all in confusion although the actual parameters are convergent in a time-based sampling process. It successfully overcomes the potential active and passive attacks in unreliable networks.

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

Besides, in the encrypted learning process to solve the complex graphical game, the evolution learning of control gain parameters is displayed by solid lines in Fig. 5, where the gain parameters in the encrypted learning converge to \( K_1 = \begin{bmatrix} -0.1753, & 0.1752 \end{bmatrix} \) and \( K_2 = \begin{bmatrix} 1.9748, & -0.0535 \end{bmatrix} \). To further illustrate the effect of solutions with quantization errors as pointed out in Remark 4, we denote the gain parameters from unprotected computing in Fig. 5 by dotted lines \( K_1^{*} = \begin{bmatrix} -0.1731, & 0.1539 \end{bmatrix} \) and \( K_2^{*} = \begin{bmatrix} 2.1768, & -0.1044 \end{bmatrix} \), which are solved from Algorithm 1 under the same initializations.

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

### TABLE I

<table>
<thead>
<tr>
<th>Time (k)</th>
<th>Actual ( \bar{P}_0 ) parameters</th>
<th>Encrypted data ( \bar{P}_0 ) computed in cloud servers with first 33 digits</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>252.161</td>
<td>760830590047951567132304367764595</td>
</tr>
<tr>
<td></td>
<td>136.513</td>
<td>63182258718176900719525328737122</td>
</tr>
<tr>
<td></td>
<td>181.167</td>
<td>58292522261300078222217638560404</td>
</tr>
<tr>
<td></td>
<td>-204.545</td>
<td>2047678085063989071770089886117</td>
</tr>
<tr>
<td>19</td>
<td>961.732</td>
<td>363054803410390184261816434470408</td>
</tr>
<tr>
<td></td>
<td>372.620</td>
<td>35010217006331600105947954640830</td>
</tr>
<tr>
<td>29</td>
<td>83.780</td>
<td>156962500098758389824626617010595</td>
</tr>
<tr>
<td></td>
<td>69.897</td>
<td>29075625565419527056366809581332</td>
</tr>
<tr>
<td></td>
<td>65.607</td>
<td>976001780709142335874828735181</td>
</tr>
<tr>
<td>39</td>
<td>82.571</td>
<td>48683953714251081320000736591560</td>
</tr>
<tr>
<td></td>
<td>36.446</td>
<td>269517829181512436708498897665</td>
</tr>
<tr>
<td></td>
<td>75.130</td>
<td>745427090107002866219416319692032</td>
</tr>
<tr>
<td>49</td>
<td>82.571</td>
<td>1480059764405452781028969782414</td>
</tr>
<tr>
<td></td>
<td>36.446</td>
<td>94248247922765597018325875025710</td>
</tr>
<tr>
<td></td>
<td>75.130</td>
<td>26812436113443840512240997205539</td>
</tr>
</tbody>
</table>
It shows that the developed encrypted learning algorithm achieves a very close solution, and guarantees the consensus against organized synchronous network attacks, as required.

For a better illustration of the consensus performance of agents with optimal policies from the graphical game, the 3-D trajectories of agents in the tracking process are presented in Fig 6, where the trajectories are projected in $[-6, 6] \times [-6, 6] \times [0, 50]$ three-dimensional space, which clearly depicts the dynamics of all five followers and the leader against synchronous network attacks.

![Fig. 6. 3-D trajectories of consensus process](image)

VI. CONCLUSION

This paper studied the global Nash equilibrium in consensus of leader-follower multi-agent dynamics against organized synchronous network attacks. A necessary and sufficient condition for the consensus and 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 in the learning process of global Nash equilibrium is guaranteed with an iteratively updated pair of decoupled gains. Based on the proposed additive-multiplicative property and Paillier’s encryption technique, an online encrypted learning algorithm has been designed, upon which 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 much encrypted data in a game.

APPENDIX

Monotonicity Proof of Theorem 3.

**Step 1:** The global performance error with $K_2^{(l)}$ can be written as

$$J_o(z, K_1^{(l)}, K_2^{(l)}) - J_o(z, K_1^*, K_2^*) = \sum_{i=1}^{N} [z^T (\bar{P}_i (K_1^{(l)}, K_2^{(l)}) z)]$$

which will be minimized by updating the gain with

$$K_1^{(l+1)} = \arg \min_{K_1} \{ \sum_{i=1}^{N} (z^T P_i (K_1^{(l)}, K_2^{(l)}) z) \}.$$  \hspace{1cm} (49)

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

$$K_1^{(l+1)} = - (\sum_{i=1}^{N} [F_i(K_1^{(l)}, K_2^{(l)})]^{-1} \sum_{i=1}^{N} [\lambda_i B^T \times P_i(K_1^{(l)}, K_2^{(l)}) (A - \lambda_i D K_2^{(l)})]$$

where $F_i(K_1^{(l)}, K_2^{(l)}) = B^T SB + \lambda_2^2 B^T P_i(K_1^{(l)}, K_2^{(l)}) B$.

**Step 2:** For the term $\bar{P}_i (K_1^{(l+1)}, K_2^{(l)})$ with any $K_1^{(l+1)}$, there is

$$J_o(z, K_1^{(l+1)}, K_2^{(l)}) - J_o(z, K_1^{(l+1)}, K_2^*) = \sum_{i=1}^{N} [z^T (\bar{P}_i (K_1^{(l+1)}, K_2^{(l)}) z)]$$

(52)

which can be minimized based on definitions of $\bar{P}_i (K_1^{(l+1)}, K_2^{(l)})$ and $P_i(K_1^{(l+1)}, K_2^{(l)})$ by

$$K_2^{(l+1)} = \arg \max_{K_2} \{ \sum_{i=1}^{N} [z^T P_i(K_1^{(l+1)}, K_2^{(l+1)}) z] \}.$$  \hspace{1cm} (53)

From computing the derivation w.r.t. disturbance input, we have

$$K_2^{(l+1)} = - (\sum_{i=1}^{N} [H_i^{(l+1)}(K_1^{(l+1)}, K_2^{(l)})]^{-1} \sum_{i=1}^{N} [\lambda_i D^T \times P_i(K_1^{(l+1)}, K_2^{(l)}) (A + \mu \lambda_i B K_1^{(l+1)})]$$

where $H_i^{(l+1)}(K_1^{(l+1)}, K_2^{(l)}) = \eta^2 D^T D - \lambda_2^2 D^T P_i(K_1^{(l+1)}, K_2^{(l)}) D$.

**Step 3:** According to Lemma 2, (50) and (54) are equal to the updating law (32). Then from the definitions of $\bar{P}_i (K_1^{(l)}, K_2^{(l)})$ and $\bar{P}_i (K_1^{(l+1)}, K_2^{(l+1)})$, we can obtain the following inequation

$$\sum_{i=1}^{N} z^T [\bar{P}_i (K_1^{(l+1)}, K_2^{(l+1)})] z$$

$$= \sum_{i=1}^{N} z^T [\bar{P}_i (K_1^{(l+1)}, K_2^{(l)}) + P_i(K_1^{(l+1)}, K_2^{(l)}) - P_i(K_1^{(l)}, K_2^{(l)})] z$$

$$= \sum_{i=1}^{N} z^T [\bar{P}_i (K_1^{(l+1)}, K_2^{(l)})] z + z^T [P_i(K_1^{(l)}, K_2^{(l)}) - P_i(K_1^{(l)}, K_2^{(l)})] z$$

$$\geq \sum_{i=1}^{N} z^T [\bar{P}_i (K_1^{(l+1)}, K_2^{(l+1)})] z$$

(56)
with \( \sum_{i=1}^{N} z_i^T [P_i(K_i^{(t+1)}, K_2^{(t+1)}) - P_i(K_i^{(t)}, K_2^{(t+1)})] z \geq 0 \), and equation

\[
P_i(K_i^{(t+1)}, K_2^{(t+1)}) = P_i(K_i^{(t)}, K_2^{(t+1)}) + P_i(K_i^{(t+1)}, K_2^{(t+1)}) - P_i(K_i^{(t)}, K_2^{(t)})
\]

(57)

\[
P_i(K_i^{(t+1)}, K_2^{(t+1)}) = P_i(K_i^{(t)}, K_2^{(t+1)}) + P_i(K_i^{(t+1)}, K_2^{(t)}) - P_i(K_i^{(t)}, K_2^{(t)})
\]

(58)

Inserting \( K_1^{(t+1)} \) and \( K_2^{(t+1)} \) into equation (33), and using (56) and (57), we obtain

\[
z_i^T P_i(l) z_i \geq z_i^T P_i(l+1) z_i \]

for all \( l \). The proof of monotonicity is thus completed. □

REFERENCES


Kun Zhang (S’18–M’20) received the B.S. degree in mathematics and applied mathematics from Hebei Normal University, China, in 2012, and the Ph.D. degree in control theory and control engineering from the Northeastern University, China, in 2020. Since 2020, he has been a postdoctoral fellow with the Institute of Systems Science, AMSS, CAS, and a research fellow with School of Electrical and Electronic Engineering, NTU. His main research interests include system optimization, adaptive control, and their industrial applications.

Ji-Feng Zhang (M’92–SM’97–F’14) received the B.S. degree in mathematics from Shandong University, China, in 1985, and the Ph.D. degree from the Institute of Systems Science (ISS), Chinese Academy of Sciences (CAS), China, in 1991. Since 1985, he has been with the ISS, CAS. His current research interests include system modeling, adaptive control, stochastic systems, and multi-agent systems. He is an IEEE Fellow, IFAC Fellow, CAA Fellow, CSIAM Fellow, member of the European Academy of Sciences and Arts, and Academician of the International Academy for Systems and Cybernetic Sciences. He received the Second Prize of the State Natural Science Award of China in 2010 and 2015, respectively. He is a Vice-President of the Chinese Mathematical Society and the Chinese Association of Automation. He was a Vice-Chair of the IFAC Technical Board, member of the Board of Governors, IEEE Control Systems Society; Convenor of Systems Science Discipline, Academic Degree Committee of the State Council of China; Vice-President of the Systems Engineering Society of China. He served as Editor-in-Chief, Deputy Editor-in-Chief, Senior Editor or Associate Editor for more than 10 journals, including Science China Information Sciences, National Science Review, IEEE Transactions on Automatic Control and SIAM Journal on Control and Optimization etc.

Huaguang Zhang (M’03–SM’04–F’14) received the B.S. and M.S. degrees in control engineering from the Northeast Dianli University of China, Jilin, China, in 1982 and 1985, respectively, and the Ph.D. degree in thermal power engineering and automation from Southeast University, Nanjing, China, in 1991. He joined the Department of Automatic Control, Northeastern University, Shenyang, China, in 1992, as a Post-Doctoral Fellow, for two years, where he has been a Professor and the Head of the Institute of Electric Automation, College of Information Science and Engineering, since 1994. He was a recipient of the Outstanding Youth Science Foundation Award from the National Natural Science Foundation Committee of China in 2003 and the IEEE Transactions on Neural Networks 2012 Outstanding Paper Award. He is an Associate Editor of Automatica, IEEE Transactions on Cybernetics, and IEEE Transactions on Neural Networks and Learning Systems etc. He is the Fellow of the E-Letter Chair of IEEE CIS Society and the former Chair of the Adaptive Dynamic Programming and Reinforcement Learning Technical Committee on IEEE Computational Intelligence Society.

Rong Su (M’11–SM’14) received the Bachelor of Engineering degree from University of Science and Technology of China in 1997, and the Master of Applied Science degree and Ph.D. degree from University of Toronto, in 2000 and 2004, respectively. He was affiliated with University of Waterloo and Technical University of Eindhoven before he joined Nanyang Technological University in 2010. Currently, he is an associate professor in the School of Electrical and Electronic Engineering. Dr. Su’s research interests include multi-agent systems, cybersecurity of discrete-event systems, supervisory control, model-based fault diagnosis, control and optimization in complex networked systems with applications in flexible manufacturing, intelligent transportation, human–robot interface, power management and green buildings. In the aforementioned areas he has more than 230 journal and conference publications, 1 monograph, and 9 granted/filed patents. Dr. Su is a senior member of IEEE, and an associate editor for Automatica, Journal of Discrete Event Dynamic Systems: Theory and Applications, and Journal of Control and Decision. He was the chair of the Technical Committee on Smart Cities in the IEEE Control Systems Society in 2016–2019, and is currently a co-chair of IEEE Robotics and Automation Society Technical Committee on Automation in Logistics, and the chair of IEEE Control Systems Chapter, Singapore. Dr Su is a recipient of 2021 Hsue-shen Tsien Paper Award from IEEE/CAA Journal of Automatica Sinica.