



Research article**Two-layer game-based consensus security control for stochastic multiagent systems over sensor attacks****Zan Li^{1,2}, Kan He^{1,2}, Zhongjun Yang^{1,2}, Lian Lian^{1,2}, Yifei Sun^{1,2} and Huaici Zhao^{1,*}**¹ College of Information Engineering, Shenyang University of Chemical Technology, Shenyang 110142, China² Key Laboratory of Information Security for Petrochemical Industry in Liaoning Province, Shenyang 110142, China*** Correspondence:** Email: hczhao@sia.cn.

Abstract: This paper investigates distributed optimal consensus security control (CSC) for uncertain multi-agent systems (MASs) with Gaussian noise, where subsystem output channels are vulnerable to sensor attacks. A novel two-layer game framework is proposed. At the agent level, the CSC problem under Gaussian processes is modeled as a zero-sum stochastic game between the controller and the attack signal, with local optimality achieved via adaptive dynamic programming (ADP). This optimality is realized by designing a distributed control protocol, treating the attack signals as independent participants and adjusting the actions based on the consensus error. At the communication network level, the Nash equilibrium of MASs under communication constraints is derived through graphical games, ensuring global optimality. The scheme guarantees that MASs maintain optimal CSC in the sense of expectation even under severe attacks, which is validated on an islanded microgrid test system.

Keywords: Adaptive dynamic programming (ADP); graphical game; multiagent systems (MASs); sensor attacks; zero-sum game

Mathematics Subject Classification: 93C10, 93A16, 93D50

1. Introduction

In recent years, the interest in multiagent systems (MASs) has significantly increased [1–4]. Given the inherent uncertainties in most control systems [5–7] and the susceptibility of the operating environment to diverse stochastic factors, such as stochastic interference and communication noise, researching stochastic uncertain MASs has gained substantial importance [8–10]. Additionally, as information exchange occurs among agents through the communication network [11], the involvement

of the communication network will lead to a profound integration of MASs with the physical systems. In particular, the communication constraints, e.g., message queues and multiple communication parameters, can significantly affect the achievable performance of multi-access and time-delay information systems [12]. This integration provides opportunities for malicious adversaries to intentionally exploit these open and shared communication networks, inevitably resulting in adverse impacts on the normal operation of MASs, and could lead to severe consequences [13]. Particularly for intricate stochastic uncertain MASs, exploring security control within unreliable communication networks has emerged as a critical focus.

Malicious attacks can manifest at various layers of cyber-physical systems, encompassing both the network and physical layers. In comparison to physical attacks, cyberattacks are more accessible due to the inherent vulnerability and unreliability of wireless networks. These attacks can obstruct the communication environment, impact control performance, and potentially result in a complete loss of control within the system [14–16]. In general, cyberattacks can be categorized into two primary types. One such security threat is jamming attacks [17], often referred to as denial-of-service (DoS) attacks. For jamming attacks, the attacker only needs to transmit jamming signals to disrupt communication between agents. Recent studies have investigated secure consensus control of MASs under DoS attacks. Wu et al. [18] developed a fault-tolerant consensus scheme with deferred actuator replacement for nonlinear MASs under DoS attacks, while Xu et al. [19] studied adaptive prescribed-performance secure consensus control for high-order constrained MASs with time-varying powers and arbitrary initial states. The other type is deception attacks [20–22], where attackers have the ability to access the physical agent itself or its components, such as sensors, controllers, and actuators, as well as communication channels, thereby being capable of manipulating the sensor measurement data and the control commands among agents. In contrast to jamming attacks that primarily disrupt data availability, deception attacks are more menacing [23]. Based on the concept of malicious attacks, the system controller and the attacker can be regarded as two players with mutually exclusive goals [24], and they attempt to select strategies that maximize their own interests and utilities by fully considering various possible competitive and confrontational behaviors of both sides. Therefore, game theory is one of the most suitable theoretical tools for characterizing such interactive behaviors and analyzing reasonable decision-making strategies.

In a competitive environment, taking advantage of the contradictory relationship between agents and uncertain attacks, the non-cooperative game naturally emerges as the primary choice for problem-solving and finds application in the field of cyber-physical systems [25, 26]. As an important branch of reinforcement learning (RL), adaptive dynamic programming (ADP) [27–29] is a powerful tool to realize optimal control within non-cooperative games. The zero-sum game is a type of non-cooperative game whose solution relies on the Hamilton-Jacobi-Isaacs (HJI) equation. However, solving HJI equations, particularly in nonlinear cases, is a difficult endeavor. Hence, investigating the security of cyber-physical systems through the employment of ADP-based zero-sum game methods is an interesting avenue of research. Jiang *et al.* [30] proposed an intelligent secure control scheme integrating optimal control theory, zero-sum game theory, and RL which transformed the secure control problem into a zero-sum game problem of an auxiliary system, leveraging the ADP method to solve the HJI equation. The proposed security control scheme can mitigate the effects of actuator attacks and perturbations. Furthermore, in order to solve the security control problem of the discrete unknown Markov jump system under false data injection attacks, Gao *et al.* [24] adopted a model-free

ADP algorithm combined with the zero-sum game control method. However, the above-mentioned cyber-physical systems are single agents, and research results are limited for more complex nonlinear MASs. Especially in complex stochastic environments, designing an optimal security control strategy for uncertain MASs based on the stochastic game will be a challenging problem.

Normally, MASs are vulnerable to various cyber attacks due to the large-scale deployment of agents and the openness of environments. However, within the existing literature, few results can be used to design the optimal security control strategy for MASs under the maximum attack condition, especially corresponding to the more complex high-order stochastic uncertain MASs. How to bring the idea of game theory into the security control of MASs has become the focus of this paper. Consequently, this paper proposes a stochastic game-based consensus security control (CSC) scheme to characterize the dynamic interaction between the uncertain MASs control input and the attacker in the multidimensional noisy environment. The primary contributions are outlined as follows.

1) A novel two-layer game framework is proposed in this paper. At the agent level, using a zero-sum stochastic game and the ADP, the proposed algorithm can find the minimum cost function under the maximum sensor attacks in the mathematical expectation sense. At the communication network level, the use of the graphical game ensures that MASs subject to communication constraints can achieve global optimality in the sense of expectation.

2) Unlike the security control methods [20–22], this paper addresses the problem of sensor attacks in a stochastic environment by proposing an optimal CSC scheme based on the zero-sum stochastic game. In this framework, the controller and the attack signal are treated as two opposing players, and the worst-case attack is explicitly considered as the saddle-point solution, so that the optimal CSC is achieved with respect to the worst attack.

3) The global Nash equilibrium of the entire MASs constrained by the communication topology is realized through a graphical game, where the local optimization of each agent is synchronized with the global optimization of the stochastic MASs. It is achieved in the sense of expectation and is well defined whenever the coupled HJI equations admit a jointly bounded admissible solution. In addition, the algorithm is applicable to MASs without stochastic noise, enhancing its universality.

Notations: In this paper, $\mathbb{R}^n$ represents real n -dimensional space. Define the set of real numbers and positive real numbers as $\mathbb{R}$ and $\mathbb{R}^+$, respectively. The trace of matrix is denoted as $\text{Tr}\{\cdot\}$. $\lambda_{\min}(\cdot)$ is the minimum eigenvalue of a matrix.

2. Preliminaries and problem formulation

2.1. Graph theory

A leader-following consensus graph is defined as $\mathcal{G}_r = (\mathcal{N}, \mathcal{E}, \mathcal{W})$, where $\mathcal{N} = \{1, \dots, N\}$ is the set of N nodes, $\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$ denotes the set of edges, and $\mathcal{W} = [a_{ij}] \in \mathbb{R}^{N \times N}$ represents the signed weight matrix of $\mathcal{G}_r$. A directed path from node i to node j is depicted by a sequence of successive edges. For the i th agent, $\mathcal{N}_i = \{j | (j, i) \in \mathcal{E}\}$ indicates its set of neighbors, where $(j, i) \in \mathcal{E}$ denotes edges from j to i and $a_{ij} \neq 0$, otherwise $a_{ij} = 0$ when $(j, i) \notin \mathcal{E}$, and supposing $a_{ii} = 0$. Define $d_i = \sum_{j=1}^N |a_{ij}|$, $\mathcal{D} = \text{diag}\{d_1, \dots, d_N\}$ and the Laplacian matrix $\mathcal{L} = \mathcal{D} - \mathcal{W}$. Throughout this paper the graph nodes are simply denoted by the agent indices i and j , and the symbol n is reserved exclusively for the state order (i.e., $\bar{x}_{i,s} = [x_{i,1}, \dots, x_{i,s}]^T \in \mathbb{R}^s$ with $s = 1, \dots, n$).

A leader node is considered and denoted as 0. In the context of the extended graph denoted as

$\bar{\mathcal{G}}_r = \{\bar{\mathcal{N}}, \bar{\mathcal{E}}\}$, $\bar{\mathcal{N}} = \{0, 1, \dots, N\}$ corresponds to the set comprising leader 0 and N follower nodes, while $\bar{\mathcal{E}} \subseteq \bar{\mathcal{N}} \times \bar{\mathcal{N}}$. b_i denotes a weighted edge from 0 to the followers. $\mathcal{B} = \text{diag}\{|b_1|, \dots, |b_N|\}$, and $H = \mathcal{L} + \mathcal{B}$.

2.2. Problem statement

Consider the model of uncertain networked MASs over unreliable communication networks, whose communication topology graph is described by $\bar{\mathcal{G}}_r$, and the i th follower agent is modeled as

$$\begin{aligned} dx_{i,s} &= (m_{i,s}x_{i,s+1} + q_{i,s}(\bar{x}_{i,s}))dt + g_{i,s}(\bar{x}_{i,s})d\varpi_i \\ dx_{i,n} &= (m_{i,n}u_i + q_{i,n}(\bar{x}_{i,n}))dt + g_{i,n}(\bar{x}_{i,n})d\varpi_i \\ y_i &= x_{i,1} + \beta_i(x_{i,1}) \end{aligned} \quad (2.1)$$

where $\bar{x}_{i,s} = [x_{i,1}, \dots, x_{i,s}]^T \in \mathbb{R}^s$ is the state vector with $s = 1, \dots, n-1$, $m_{i,s}$ denotes a positive dynamic parameter, $q_{i,s}(\bar{x}_{i,s})$ and $g_{i,s}(\bar{x}_{i,s})$ represent the unknown smooth nonlinear functions, u_i denotes the control input, and y_i is the measurement output. ϖ_i signifies the independent r_i -dimensional standard Wiener process.

Attackers usually steal system information to optimize attack functions and reduce system performance. According to the result presented in [31], it is assumed that attackers employ more versatile state nonlinear functions within the attack information, and thus can pretend to be included in the system information to confuse the detector. The sensor attacks over unreliable communication networks are represented by

$$\beta_i(x_{i,1}) = \varphi_i(x_{i,1})p_i + k_i \quad (2.2)$$

where p_i denotes the unknown parameter, $\varphi_i(x_{i,1})$ is the known regression function, and k_i represents a regression error, which are the zero-mean independent identically distributed Gaussian noises.

The control purpose is according to the stochastic game-based CSC scheme designed to realize the uniformly ultimately bounded (UUB) of all errors in the closed-loop system under the condition of minimizing the cost function.

To facilitate subsequent design, the following assumptions are required.

Assumption 1. [10, 32] The leader's trajectory function $y_r(t) \in \mathbb{R}$ and its derivatives $\dot{y}_r(t) \in \mathbb{R}$ are continuous, which can be acquired by followers if $b_i \neq 0$ with $i = 1, \dots, N$.

Assumption 2. [32] The positive dynamic parameter $m_{i,s}(\cdot)$ is a known smooth function (a physical coefficient of the i th subsystem) which is uniformly bounded away from zero. There exist positive constants $m_{\min}^{i,s}$ and $m_{\max}^{i,s}$ such that the function $m_{i,s}(\cdot)$ is bounded, satisfying $m_{\min}^{i,s} \leq |m_{i,s}(\cdot)| \leq m_{\max}^{i,s}$.

Definition 1. [32] $V(x, t)$ denotes a function that possesses two continuous derivatives. $q(x, t)$ and $g(x, t)$ are unknown smooth nonlinear functions. Define the differential operator $\mathcal{L}$ and the derivative d as follows

$$\mathcal{L}V = \frac{\partial V}{\partial t} + \frac{\partial V}{\partial x}q + \frac{1}{2}\text{Tr}\{g^T \frac{\partial^2 V}{\partial x^2} g\} \quad (2.3)$$

$$dV = \mathcal{L}Vdt + \frac{\partial V}{\partial x}gd\varpi. \quad (2.4)$$

Lemma 1. [33] A continuous function $\Gamma(x)$ is defined over the compact set Ξ . Then, for a given desired level of accuracy $\epsilon > 0$, there exists the fuzzy logic systems (FLSs) $\xi^T \Phi(x)$ such that

$$\sup_{x \in \Xi} |\Gamma(x) - \xi^T \Phi(x)| \leq \epsilon.$$

$\xi = [\xi_1, \xi_2, \dots, \xi_\kappa]^T$ denotes the ideal constant weight vector, where κ represents the number of fuzzy rules, $\Phi(x) = [\phi_1(x), \phi_2(x), \dots, \phi_\kappa(x)]^T / \sum_{r=1}^{\kappa} \phi_r(x)$ denotes a basis function vector, and ξ^* is the optimal parameter vector. Then

$$\xi^* := \arg \min_{\xi \in \mathbb{R}^\kappa} \{ \sup_{x \in \Xi} |\Gamma(x) - \xi^T \Phi(x)| \}.$$

The minimum fuzzy approximation error is given

$$\epsilon = \Gamma(x) - \xi^{*T} \Phi(x).$$

3. Two-layer game-based consensus security control design

A two-layer game framework is designed in this section. First, in the output feedback layer of MASs, a new coordinate transformation is constructed by backstepping technology, so that MASs under sensor attacks can still achieve consensus control. Subsequently, the consensus errors dynamics model is established. In the consensus error feedback layer, the stochastic game is used to describe the interaction process between the controllers and the malicious attackers. Then, the CSC problem of stochastic MASs is transformed into the equivalent stochastic game problem of consensus error feedback dynamics. In order to illustrate the proposed control method in more detail, the structure diagram of the two-layer game framework is shown in Figure 1.

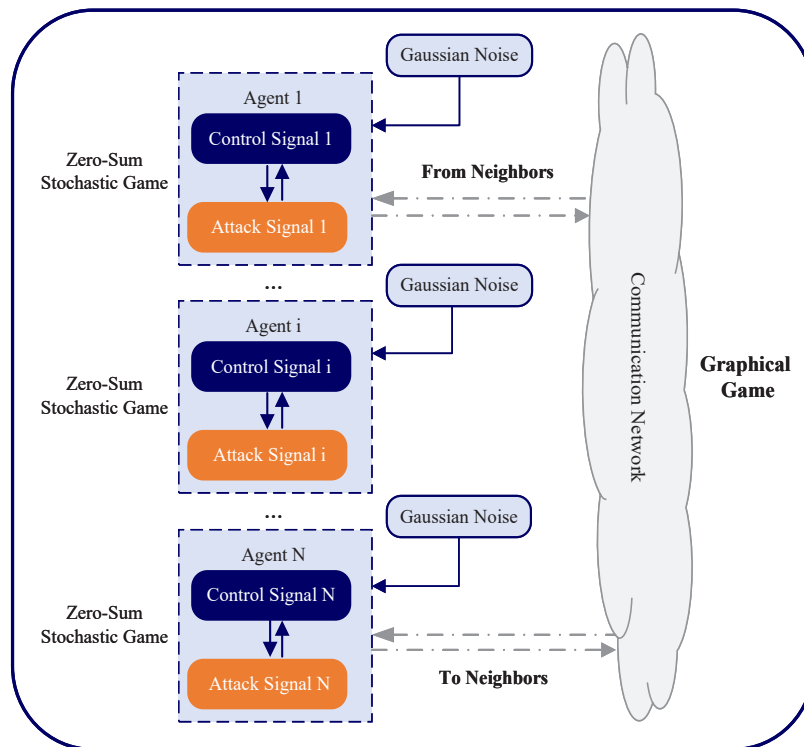


Figure 1. The structure diagram of the two-layer framework.

3.1. Adaptive consensus security protocol design

This section focuses on the design of an adaptive consensus security protocol for stochastic MASs with sensor attacks over unreliable communication networks. To facilitate the separation of sensor attack signals to construct the game strategy, a set of dynamic variables $z_{i,1}$ ($i = 1, 2, \dots, N$) is designed, and the following error transformation is constructed:

$$\begin{cases} e_{i,1} = y_i - z_{i,1} \\ e_{i,s} = x_{i,s} - \bar{\alpha}_{i,s} \\ \dot{z}_{i,1} = - \sum_{j=1}^N a_{ij}(y_i - y_j) - b_i(y_i - y_r) \end{cases} \quad (3.1)$$

where $s = 2, 3, \dots, n$. $z_{i,1}$ signifies the difference in outputs between the i th agent and its neighbors, with an initial condition of $z_{i,1}(0) = x_{i,1}(0)$. Additionally, the virtual controllers $\bar{\alpha}_{i,s}$ are derived $x_{i,sr}$ via the first-order command filter.

The virtual controller $x_{i,sr}$ ($s = 2, \dots, n + 1$) and the actual control $x_{i,nr} = u_i$ are defined as

$$x_{i,sr} = x_{i,sr}^\alpha + x_{i,sr}^* \quad (3.2)$$

where $x_{i,sr}^\alpha$ and $x_{i,(n+1)r}^\alpha = u_i^\alpha$ represent virtual input and actual control input, respectively. $x_{i,sr}^*$ denotes the optimal control input.

Define the following first-order filter:

$$\tau_{i,s} \dot{\bar{\alpha}}_{i,s} + \bar{\alpha}_{i,s} = x_{i,sr}^\alpha, \bar{\alpha}_{i,s}(0) = x_{i,sr}(0) \quad (3.3)$$

where $\tau_{i,s}$ is a constant with $s = 2, \dots, n$. Then, the error transformation is constructed:

$$z_{i,s} = \bar{\alpha}_{i,s} - x_{i,sr} \quad (3.4)$$

with $s = 2, \dots, n$.

The backstepping designed process is outlined below. Define $\theta_{i,s} = \|\xi_{i,s}^*\|^2$ ($s = 1, \dots, n$). Let $\hat{\theta}_{i,s}$ be an estimate of $\theta_{i,s}$, and the estimation error is $\tilde{\theta}_{i,s} = \theta_{i,s} - \hat{\theta}_{i,s}$.

Step 1 : Choose a Lyapunov function candidate

$$V_{i,1} = \frac{1}{4}e_{i,1}^4 + \frac{1}{2\varrho_{i,1}}\tilde{\theta}_{i,1}^2 + \frac{1}{4}z_{i,2}^4 \quad (3.5)$$

where $\varrho_{i,1}$ is a positive constant.

Using (3.1) and the Itô formula, there is

$$\begin{aligned} \mathcal{L}V_{i,1} = & e_{i,1}^3 \left[m_{i,1}(e_{i,2} + \bar{\alpha}_{i,2}) + \dot{\beta}_i + q_{i,1}(x_{i,1}) - b_i y_r + (d_i + b_i)(x_{i,1} + \beta_i) - \sum_{j=1}^N a_{ij}(x_{j,1} + \beta_j) \right] \\ & - \frac{\tilde{\theta}_{i,1}\dot{\hat{\theta}}_{i,1}}{\varrho_{i,1}} + \frac{3}{2}e_{i,1}^2 g_{i,1}^T g_{i,1} + z_{i,2}^3 \dot{z}_{i,2}. \end{aligned} \quad (3.6)$$

Applying Young's inequality, and $l_{i,1} > 0$, the subsequent inequalities are valid:

$$\frac{3}{2}e_{i,1}^2 g_{i,1}^T g_{i,1} \leq \frac{3}{4}l_{i,1}^{-2}e_{i,1}^4 \|g_{i,1}\|^4 + \frac{3}{4}l_{i,1}^2 \quad (3.7)$$

$$m_{i,1}e_{i,1}^3 e_{i,2} \leq \frac{3m_{i,1}}{4}e_{i,1}^4 + \frac{m_{i,1}}{4}e_{i,2}^4. \quad (3.8)$$

Substituting (3.7) and (3.8) into (3.6) yields

$$\mathcal{L}V_{i,1} \leq e_{i,1}^3 \left[m_{i,1}\bar{\alpha}_{i,2} + \Gamma_{i,1}(T_{i,1}) + (d_i + b_i)\beta_i - b_i y_r \right] - \frac{\tilde{\theta}_{i,1}\dot{\hat{\theta}}_{i,1}}{\varrho_{i,1}} + z_{i,2}^3 \dot{z}_{i,2} + \frac{3}{4}l_{i,1}^2 + \frac{m_{i,1}}{4}e_{i,2}^4 + e_{i,1}^3 h_{i,1}(e_{i,1}) \quad (3.9)$$

where

$$\Gamma_{i,1}(T_{i,1}) = q_{i,1}(x_{i,1}) + \dot{\beta}_i - (d_i + b_i)x_{i,1} + \frac{3}{4}l_{i,1}^{-2}e_{i,1}\|g_{i,1}\|^4 - \sum_{j=1}^N a_{ij}(x_{j,1} + \beta_j) + \frac{3m_{i,1}}{4}e_{i,1} - h_{i,1}(e_{i,1}) \quad (3.10)$$

with $T_{i,1} = [x_{i,1}, x_{j,1}, \dot{\beta}_i, \beta_j]^T$, and $h_{i,1}(e_{i,1})$ is a smooth function with respect to $e_{i,1}$.

Using the error transformation (3.4), one has

$$\mathcal{L}V_{i,1} \leq e_{i,1}^3 \left[m_{i,1}(z_{i,2} + x_{i,2r}) + h_{i,1}(e_{i,1}) + (d_i + b_i)\beta_i - b_i y_r + \Gamma_{i,1}(T_{i,1}) \right] - \frac{\tilde{\theta}_{i,1}\dot{\hat{\theta}}_{i,1}}{\varrho_{i,1}} + z_{i,2}^3 \dot{z}_{i,2} + \frac{3}{4}l_{i,1}^2 + \frac{m_{i,1}}{4}e_{i,2}^4 \quad (3.11)$$

The FLSs are utilized to approximate the unknown term $\Gamma_{i,s}(T_{i,s})$ ($s = 1, \dots, n$), and one has

$$\Gamma_{i,s}(T_{i,s}) = \xi_{i,s}^{*T} \Phi_{i,s}(T_{i,s}) + \epsilon_{i,s} \quad (3.12)$$

where $\epsilon_{i,s}$ is the minimum approximation error. Here the FLSs approximate the compound uncertainties $\Gamma_{i,s}(T_{i,s})$, which collect the unknown dynamics $q_{i,s}$, the smooth functions $h_{i,s}$, and the Itô-noise related terms.

Applying Young's inequality and $a_{i,1} > 0$, there are

$$e_{i,1}^3 \Gamma_{i,1}(T_{i,1}) \leq \frac{1}{2a_{i,1}^2} \theta_{i,1} e_{i,1}^6 \Phi_{i,1}^T(T_{i,1}) \Phi_{i,1}(T_{i,1}) + \frac{3}{4}e_{i,1}^4 + \frac{a_{i,1}^2}{2} + \frac{\epsilon_{i,1}^4}{2} \quad (3.13)$$

$$m_{i,1}e_{i,1}^3 z_{i,2} \leq \frac{3}{4}e_{i,1}^4 + \frac{m_{i,1}^4}{4}z_{i,2}^4. \quad (3.14)$$

Substituting (3.2) and (3.12)–(3.14) into (3.11) yields

$$\begin{aligned} \mathcal{L}V_{i,1} \leq & e_{i,1}^3 \left[h_{i,1}(e_{i,1}) + (d_i + b_i)\beta_i - b_i y_r + m_{i,1}(x_{i,2r}^\alpha + x_{i,2r}^*) \right] - \frac{\tilde{\theta}_{i,1}\dot{\hat{\theta}}_{i,1}}{\varrho_{i,1}} + \frac{\epsilon_{i,s}^4}{2} + z_{i,2}^3 \dot{z}_{i,2} \\ & + \frac{m_{i,1}^4}{4}z_{i,2}^4 + \frac{m_{i,1}}{4}e_{i,2}^4 + \frac{a_{i,1}^2}{2} + \frac{3}{2}e_{i,1}^4 + \frac{1}{2a_{i,1}^2} \theta_{i,1} e_{i,1}^6 \Phi_{i,1}^T(T_{i,1}) \Phi_{i,1}(T_{i,1}) + \frac{3}{4}l_{i,1}^2. \end{aligned} \quad (3.15)$$

Design a virtual controller and an adaptive law as

$$x_{i,2r}^\alpha = \frac{1}{m_{i,1}} \left[-c_{i,1}e_{i,1} - \frac{3}{2}e_{i,1} + b_i y_r - \frac{1}{2a_{i,1}^2} \hat{\theta}_{i,1} e_{i,1}^3 \Phi_{i,1}^T(T_{i,1}) \Phi_{i,1}(T_{i,1}) \right] \quad (3.16)$$

$$\dot{\hat{\theta}}_{i,1} = \frac{\varrho_{i,1}}{2a_{i,1}^2} e_{i,1}^6 \Phi_{i,1}^T(T_{i,1}) \Phi_{i,1}(T_{i,1}) - \gamma_{i,1} \hat{\theta}_{i,1} \quad (3.17)$$

where $c_{i,1}$ and $\gamma_{i,1}$ denote positive constants.

By substituting (3.16)–(3.17) into (3.15), we obtain

$$\begin{aligned} \mathcal{L}V_{i,1} \leq & e_{i,1}^3 \left[h_{i,1}(e_{i,1}) + (d_i + b_i)\beta_i + m_{i,1}x_{i,1r}^* \right] + \frac{m_{i,1}^4}{4} z_{i,2}^4 + \frac{3}{4} l_{i,1}^2 - c_{i,1}e_{i,2}^4 + \frac{m_{i,1}}{4} e_{i,2}^4 \\ & + \frac{\gamma_{i,1} \tilde{\theta}_{i,1} \hat{\theta}_{i,1}}{\varrho_{i,1}} + \frac{a_{i,1}^2}{2} + \frac{\epsilon_{i,1}^4}{2} + z_{i,2}^3 \dot{z}_{i,2}. \end{aligned} \quad (3.18)$$

Step s ($2 \leq s \leq n-1$): Choose the following Lyapunov function:

$$V_{i,s} = V_{i,s-1} + \frac{1}{4} e_{i,s}^4 + \frac{1}{2\varrho_{i,s}} \tilde{\theta}_{i,s}^2 + \frac{1}{4} z_{i,s+1}^4 \quad (3.19)$$

where $\varrho_{i,s}$ is a positive constant.

According to the Itô formula, (3.1), (3.3), and (3.4), one has

$$\begin{aligned} \mathcal{L}V_{i,s} = & \mathcal{L}V_{i,s-1} - \frac{\tilde{\theta}_{i,s} \dot{\tilde{\theta}}_{i,s}}{\varrho_{i,s}} + z_{i,s+1}^3 \dot{z}_{i,s+1} + \frac{3}{2} e_{i,s}^2 g_{i,s}^T g_{i,s} \\ & + e_{i,s}^3 \left[m_{i,s}(e_{i,s+1} + z_{i,s+1} + x_{i,(s+1)r}) + q_{i,s}(\bar{x}_{i,s}) - \frac{x_{i,sr}^\alpha - \bar{\alpha}_{i,s}}{\tau_{i,s}} \right]. \end{aligned} \quad (3.20)$$

Applying Young's inequality, for $l_{i,s} > 0$, there are

$$\frac{3}{2} e_{i,s}^2 g_{i,s}^T g_{i,s} \leq \frac{3}{4} l_{i,s}^{-2} e_{i,s}^4 \|g_{i,s}\|^4 + \frac{3}{4} l_{i,s}^2 \quad (3.21)$$

$$m_{i,s} e_{i,s}^3 e_{i,s+1} \leq \frac{3m_{i,s}}{4} e_{i,s}^4 + \frac{m_{i,s}}{4} e_{i,s+1}^4 \quad (3.22)$$

$$m_{i,s} e_{i,s}^3 z_{i,s+1} \leq \frac{3}{4} e_{i,s}^4 + \frac{m_{i,s}^4}{4} z_{i,s+1}^4. \quad (3.23)$$

Substituting (3.21)–(3.23) into (3.20) yields

$$\begin{aligned} \mathcal{L}V_{i,s} \leq & e_{i,1}^3 (d_i + b_i)\beta_i - \frac{\tilde{\theta}_{i,s} \dot{\tilde{\theta}}_{i,s}}{\varrho_{i,s}} + \frac{m_{i,s}}{4} e_{i,s+1}^4 + \frac{3}{4} l_{i,s}^2 + \sum_{p=1}^{s-1} e_{i,p}^3 [h_{i,p}(e_{i,p}) + m_{i,p} x_{i,pr}^*] \\ & + \sum_{p=1}^{s-1} \left(\frac{3}{4} l_{i,p}^2 - c_{i,p} e_{i,p}^4 + \frac{m_{i,p}^4}{4} z_{i,p+1}^4 + \frac{a_{i,p}^2}{2} + \frac{\gamma_{i,p} \tilde{\theta}_{i,p} \hat{\theta}_{i,p}}{\varrho_{i,p}} + z_{i,p+1}^3 \dot{z}_{i,p+1} + \frac{\epsilon_{i,p}^4}{2} \right) \\ & + e_{i,s}^3 \left[\Gamma_{i,s}(T_{i,s}) + m_{i,s} x_{i,sr} - \frac{x_{i,sr}^\alpha - \bar{\alpha}_{i,s}}{\tau_{i,s}} \right] + z_{i,s+1}^3 \dot{z}_{i,s+1} + \frac{m_{i,s}^4}{4} z_{i,s+1}^4 + e_{i,s}^3 h_{i,s}(e_{i,s}) \end{aligned} \quad (3.24)$$

where $\Gamma_{i,s}(T_{i,s}) = q_{i,s}(\bar{x}_{i,s}) + \frac{3}{4}l_{i,s}^2 e_{i,s} \|g_{i,s}\|^4 - h_{i,s}(e_{i,s}) + \frac{m_{i,s-1} + 3m_{i,s} + 3}{4} e_{i,s}$ with $T_{i,s} = \bar{x}_{i,s}$, and $h_{i,s}(e_{i,s})$ is a smooth function with respect to $e_{i,s}$.

Similar to the Step 1, utilizing (3.12) and Young's inequality, for $a_{i,s} > 0$, (3.24) is derived as

$$\begin{aligned} \mathcal{L}V_{i,s} \leq & e_{i,1}^3(d_i + b_i)\beta_i - \frac{\tilde{\theta}_{i,s}\hat{\theta}_{i,s}}{\varrho_{i,s}} + \frac{m_{i,s}}{4}e_{i,s+1}^4 + \frac{\epsilon_{i,s}^4}{2} + \sum_{p=1}^{s-1} e_{i,p}^3[h_{i,p}(e_{i,p}) + m_{i,p}x_{i,pr}^*] + z_{i,s+1}^3\dot{z}_{i,s+1} \\ & + \sum_{p=1}^{s-1} \left(\frac{3}{4}l_{i,p}^2 - c_{i,p}e_{i,p}^4 + \frac{m_{i,p}^4}{4}z_{i,p+1}^4 + \frac{a_{i,p}^2}{2} + \frac{\gamma_{i,p}\tilde{\theta}_{i,p}\hat{\theta}_{i,p}}{\varrho_{i,p}} + z_{i,p+1}^3\dot{z}_{i,p+1} + \frac{\epsilon_{i,p}^4}{2} \right) + \frac{a_{i,s}^2}{2} \\ & + e_{i,s}^3 \left[h_{i,s}(e_{i,s}) + m_{i,s}(x_{i,(s+1)r}^\alpha + x_{i,(s+1)r}^*) - \frac{x_{i,sr}^\alpha - \bar{\alpha}_{i,s}}{\tau_{i,s}} \right] + \frac{3}{4}e_{i,s}^4 + \frac{3}{4}l_{i,s}^2 + \frac{m_{i,s}^4}{4}z_{i,s+1}^4 \\ & + \frac{1}{2a_{i,s}^2}\theta_{i,s}e_{i,s}^6\Phi_{i,s}^T(T_{i,s})\Phi_{i,s}(T_{i,s}). \end{aligned} \quad (3.25)$$

Design a virtual controller and an adaptive law

$$x_{i,(s+1)r}^\alpha = \frac{1}{m_{i,s}} \left[-c_{i,s}e_{i,s} - \frac{3}{4}e_{i,s} + \frac{x_{i,sr}^\alpha - \bar{\alpha}_{i,s}}{\tau_{i,s}} - \frac{\hat{\theta}_{i,s}}{2a_{i,s}^2}e_{i,s}^3\Phi_{i,s}^T(T_{i,s})\Phi_{i,s}(T_{i,s}) \right] \quad (3.26)$$

$$\hat{\theta}_{i,s} = \frac{\varrho_{i,s}}{2a_{i,s}^2}e_{i,s}^6\Phi_{i,s}^T(T_{i,s})\Phi_{i,s}(T_{i,s}) - \gamma_{i,s}\hat{\theta}_{i,s} \quad (3.27)$$

where $c_{i,s}$ and $\varrho_{i,s}$ are the positive design constants.

Substituting (3.26)–(3.27) into (3.25) yields

$$\begin{aligned} \mathcal{L}V_{i,s} \leq & \sum_{p=1}^s \left(z_{i,p+1}^3\dot{z}_{i,p+1} - c_{i,p}e_{i,p}^4 + \frac{m_{i,p}^4}{4}z_{i,p+1}^4 + \frac{3}{4}l_{i,p}^2 - \frac{\gamma_{i,p}\tilde{\theta}_{i,p}\hat{\theta}_{i,p}}{\varrho_{i,p}} + \frac{a_{i,p}^2}{2} + \frac{\epsilon_{i,p}^4}{2} \right) \\ & + e_{i,1}^3(d_i + b_i)\beta_i + \frac{m_{i,s}}{4}e_{i,s+1}^4 + \sum_{p=1}^s e_{i,p}^3[h_{i,p}(e_{i,p}) + m_{i,p}x_{i,pr}^*]. \end{aligned} \quad (3.28)$$

Step n : Choose a Lyapunov function candidate

$$V_{i,n} = V_{i,n-1} + \frac{1}{4}e_{i,n}^4 + \frac{1}{2\varrho_{i,n}}\tilde{\theta}_{i,n}^2 \quad (3.29)$$

where $\varrho_{i,n}$ is a positive constant.

According to Itô formula and based on (3.1) and (3.3), one has

$$\mathcal{L}V_{i,n} = \mathcal{L}V_{i,n-1} + \frac{3}{2}e_{i,n}^2g_{i,n}^Tg_{i,n} - \frac{\tilde{\theta}_{i,n}\hat{\theta}_{i,n}}{\varrho_{i,n}} + e_{i,n}^3 \left[m_{i,n}u_i + q_{i,n}(\bar{x}_{i,n}) - \frac{x_{i,nr}^\alpha - \bar{\alpha}_{i,n}}{\tau_{i,n}} \right]. \quad (3.30)$$

Using Young's inequality, for $l_{i,n} > 0$, we obtain

$$\mathcal{L}V_{i,n} \leq e_{i,1}^3(d_i + b_i)\beta_i + \sum_{p=1}^{n-1} e_{i,p}^3[h_{i,p}(e_{i,p}) + m_{i,p}x_{i,pr}^*] - \frac{\tilde{\theta}_{i,n}\hat{\theta}_{i,n}}{\varrho_{i,n}} + \frac{3}{4}l_{i,n}^2 + e_{i,n}^3h_{i,n}(e_{i,n})$$

$$\begin{aligned}
& + \sum_{p=1}^{n-1} \left(\frac{3}{4} l_{i,p}^2 - c_{i,p} e_{i,p}^4 + \frac{m_{i,p}^4}{4} z_{i,p+1}^4 + \frac{a_{i,p}^2}{2} - \frac{\gamma_{i,p} \tilde{\theta}_{i,p} \hat{\theta}_{i,p}}{\varrho_{i,p}} + z_{i,p+1}^3 \dot{z}_{i,p+1} + \frac{\epsilon_{i,p}^4}{2} \right) \\
& + e_{i,n}^3 \left[m_{i,n} u_i + \Gamma_{i,n}(T_{i,n}) - \frac{x_{i,nr}^\alpha - \bar{\alpha}_{i,n}}{\tau_{i,n}} \right]
\end{aligned} \quad (3.31)$$

where

$$\Gamma_{i,n}(T_{i,n}) = q_{i,n}(\bar{x}_{i,n}) + \frac{3}{4} l_{i,n}^2 e_{i,n} \|g_{i,n}\|^4 - h_{i,n}(e_{i,n} + \frac{m_{i,n-1}}{4} e_{i,n})$$

with $T_{i,n} = \bar{x}_{i,n}$.

According to (3.12) and Young's inequality, for $a_{i,n} > 0$, (3.31) is derived as

$$\begin{aligned}
\mathcal{L}V_{i,n} & \leq e_{i,1}^3 (d_i + b_i) \beta_i + \frac{3}{4} l_{i,n}^2 + \frac{\epsilon_{i,n}^4}{2} + \frac{a_{i,n}^2}{2} + \frac{3}{4} e_{i,n}^4 + \sum_{p=1}^{n-1} e_{i,p}^3 \left[h_{i,p}(e_{i,p}) + m_{i,p} x_{i,pr}^* \right] \\
& + \sum_{p=1}^{n-1} \left(\frac{3}{4} l_{i,p}^2 - c_{i,p} e_{i,p}^4 + \frac{m_{i,p}^4}{4} z_{i,p+1}^4 + \frac{a_{i,p}^2}{2} - \frac{\gamma_{i,p} \tilde{\theta}_{i,p} \hat{\theta}_{i,p}}{\varrho_{i,p}} + z_{i,p+1}^3 \dot{z}_{i,p+1} + \frac{\epsilon_{i,p}^4}{2} \right) - \frac{\tilde{\theta}_{i,n} \hat{\theta}_{i,n}}{\varrho_{i,n}} \\
& + e_{i,n}^3 \left[h_{i,n}(e_{i,n}) + m_{i,n} (u_i^\alpha + x_{i,nr}^*) - \frac{x_{i,nr}^\alpha - \bar{\alpha}_{i,n}}{\tau_{i,n}} \right] + \frac{1}{2a_{i,n}^2} \theta_{i,n} e_{i,n}^6 \Phi_{i,n}^T(T_{i,n}) \Phi_{i,n}(T_{i,n}).
\end{aligned} \quad (3.32)$$

The actual controller and the adaptive law are constructed as

$$u_i^\alpha = \frac{1}{m_{i,n}} \left[-c_{i,n} e_{i,n} - \frac{3}{4} e_{i,n} + \frac{x_{i,nr}^\alpha - \bar{\alpha}_{i,n}}{\tau_{i,n}} - \frac{\hat{\theta}_{i,n}}{2a_{i,n}^2} e_{i,n}^3 \Phi_{i,n}^T(T_{i,n}) \Phi_{i,n}(T_{i,n}) \right] \quad (3.33)$$

$$\dot{\hat{\theta}}_{i,n} = \frac{\varrho_{i,n}}{2a_{i,n}^2} e_{i,n}^6 \Phi_{i,n}^T(T_{i,n}) \Phi_{i,n}(T_{i,n}) - \gamma_{i,n} \hat{\theta}_{i,n} \quad (3.34)$$

where $c_{i,n}$ and $\varrho_{i,n}$ are the positive design constants.

Substituting (3.33)–(3.34) into (3.32) yields

$$\begin{aligned}
\mathcal{L}V_{i,n} & \leq e_{i,1}^3 (d_i + b_i) \beta_i + \sum_{p=1}^n \left(\frac{\gamma_{i,p} \tilde{\theta}_{i,p} \hat{\theta}_{i,p}}{\varrho_{i,p}} + \frac{3}{4} l_{i,p}^2 + \frac{a_{i,p}^2}{2} + \frac{\epsilon_{i,p}^4}{2} - c_{i,p} e_{i,p}^4 \right) \\
& + \sum_{p=1}^{n-1} \left(\frac{m_{i,p}^4}{4} z_{i,p+1}^4 + z_{i,p+1}^3 \dot{z}_{i,p+1} \right) + \sum_{p=1}^n e_{i,p}^3 \left[h_{i,p}(e_{i,p}) + m_{i,p} x_{i,pr}^* \right].
\end{aligned} \quad (3.35)$$

It follows from $\tau_{i,s} \dot{\bar{\alpha}}_{i,s} + \bar{\alpha}_{i,s} = x_{i,sr}^\alpha$ that

$$\dot{z}_{i,s} = -\frac{z_{i,s}}{\tau_{i,s}} + B_{i,s}(\cdot), \quad s = 2, \dots, n$$

where

$$B_{i,2}(\cdot) = c_{i,1} \dot{e}_{i,1} + \frac{\dot{\hat{\theta}}_{i,1}}{2a_{i,1}} \left[e_{i,1}^3 \Phi_{i,1}^T(T_{i,1}) \Phi_{i,1}(T_{i,1}) \right] - b_i \dot{y}_r + \frac{3}{4} \dot{e}_{i,1}$$

$$+ \frac{3\hat{\theta}_{i,1}}{2a_{i,1}} \left[e_{i,1}^2 \dot{e}_{i,1} \Phi_{i,1}^T(T_{i,1}) \Phi_{i,1}(T_{i,1}) \right] + \frac{\hat{\theta}_{i,1}}{a_{i,1}} \left[e_{i,1}^3 \Phi_{i,1}^T(T_{i,1}) \dot{\Phi}_{i,1}(T_{i,1}) \right]$$

and for $s = 3, \dots, n$

$$B_{i,s+1}(\cdot) = c_{i,s} \dot{e}_{i,s} + \frac{\dot{\hat{\theta}}_{i,s}}{2a_{i,s}} \left[e_{i,s}^3 \Phi_{i,s}^T(T_{i,s}) \Phi_{i,s}(T_{i,s}) \right] - \frac{\dot{z}_{i,s}}{\tau_{i,s}} \\ + \frac{\hat{\theta}_{i,s}}{a_{i,s}} \left[e_{i,s}^3 \Phi_{i,s}^T(T_{i,s}) \dot{\Phi}_{i,s}(T_{i,s}) \right] + \frac{3\hat{\theta}_{i,s}}{2a_{i,s}} \left[e_{i,s}^2 \dot{e}_{i,s} \Phi_{i,s}^T(T_{i,s}) \Phi_{i,s}(T_{i,s}) \right].$$

Remark 1. It is emphasized that $B_{i,p+1}(\cdot)$ is not a prescribed constant bound, but an explicitly constructed function of the known signals $e_{i,s}$ and $\hat{\theta}_{i,s}$ and the fuzzy basis functions, whose exact form is given by $B_{i,2}(\cdot)$ and $B_{i,s+1}(\cdot)$ above. Since $B_{i,p+1}(\cdot)$ is composed of continuous functions of bounded closed-loop signals and bounded basis functions, it is bounded on the compact set under consideration, i.e., $|B_{i,p+1}(\cdot)| \leq \bar{B}_{i,p+1}$. Consequently, the inequality required by the closed-loop analysis, namely

$$\bar{\tau}_{i,p+1} = \frac{1}{\tau_{i,p}} - \frac{m_{i,p}^4}{4} - \frac{3}{4}(\kappa_{i,p} \bar{B}_{i,p+1})^{\frac{4}{3}} > 0,$$

is guaranteed in practical tuning by selecting the filter time constant $\tau_{i,p}$ sufficiently small and the Young's inequality parameter $\kappa_{i,p}$ appropriately. This tuning rule is used in the simulation studies.

Using Young's inequality and $\tilde{\theta}_{i,p} = \theta_{i,p} - \hat{\theta}_{i,p}$, (3.35) is further calculated as

$$\mathcal{L}V_{i,n} \leq e_{i,1}^3 (d_i + b_i) \beta_i + \sum_{p=1}^n e_{i,p}^3 \left[h_{i,p}(e_{i,p}) + m_{i,p} x_{i,pr}^* \right] \\ + \sum_{p=1}^{n-1} \left(-\frac{z_{i,p+1}^4}{\tau_{i,p}} + \frac{3}{4}(\kappa_{i,p} B_{i,p+1})^{\frac{4}{3}} z_{i,p+1}^4 + \frac{m_{i,p}^4}{4} z_{i,p+1}^4 + \frac{1}{4\kappa_{i,p}^4} \right) \\ + \sum_{p=1}^n \left(-c_{i,p} e_{i,p}^4 - \frac{\gamma_{i,p} \|\tilde{\theta}_{i,p}\|^2}{2Q_{i,p}} + \frac{\gamma_{i,p} \|\theta_{i,p}\|^2}{2Q_{i,p}} + \frac{a_{i,p}^2}{2} + \frac{\epsilon_{i,p}^4}{2} + \frac{3}{4} l_{i,p}^2 \right) \quad (3.36)$$

where $\kappa_{i,p}$ is the positive constant.

For all agents, considering the Lyapunov function $V_S = \sum_{i=1}^N V_{i,n}$

$$\mathcal{L}V_S \leq \sum_{i=1}^N \left(\sum_{p=1}^n e_{i,p}^3 \left[h_{i,p}(e_{i,p}) + m_{i,p} x_{i,pr}^* \right] + e_{i,1}^3 (d_i + b_i) \beta_i \right) - \sum_{i=1}^N \sum_{p=1}^n c_{i,p} e_{i,p}^4 \\ + \sum_{i=1}^N \sum_{p=1}^{n-1} z_{i,p+1}^4 \left(\frac{3}{4}(\kappa_{i,p} B_{i,p+1})^{\frac{4}{3}} - \frac{1}{\tau_{i,p}} + \frac{m_{i,p}^4}{4} \right) + \sum_{i=1}^N P_i - \sum_{i=1}^N \sum_{p=1}^n \frac{\gamma_{i,p} \|\tilde{\theta}_{i,p}\|^2}{2Q_{i,p}} \quad (3.37)$$

where

$$P_i = \sum_{p=1}^n \left(\frac{\gamma_{i,p} \|\theta_{i,p}\|^2}{2Q_{i,p}} + \frac{a_{i,p}^2}{2} + \frac{\epsilon_{i,p}^4}{2} + \frac{3}{4} l_{i,p}^2 \right) + \sum_{p=1}^{n-1} \frac{1}{4\kappa_{i,p}^4}.$$

In what follows we absorb the bounded Itô second-order term of the stochastic game layer into a unified constant. Since $\frac{1}{2}\text{Tr}\{G_i^T \nabla^2 J_{si}(E_i) G_i\} \leq \bar{P}_i$ (because both G_i and $\nabla^2 J_{si}(E_i)$ are bounded on the compact set), we define the combined residual constant

$$P_i^s := P_i + \bar{P}_i,$$

which is used in place of P_i in all the ultimate bounds derived below.

Remark 2. For uncertain stochastic MASs under sensor attacks, the main difficulty of leader-follower consensus control is that none of the agents can receive the correct leader information $y_r(t)$ through the sensor, and the uncertainty of the nonlinear sensor attacks and the state of MASs are intertwined, making the problem more complicated. The key to solving the above problem lies in the design of coordinate transformation (3.1) so that the sensor attack is easy to be separated and used for the design of the stochastic game.

3.2. Stochastic game-based control design

In this section, considering the consensus tracking error dynamics over sensor attacks, stochastic game theory is utilized, and the Nash equilibrium solution (U_i^*, Υ_i^*) with $U_i^* = [x_{i,1r}^*, \dots, x_{i,nr}^*]^T \in \mathbb{R}^n$ and $\Upsilon_i^* = [\beta_i^*, 1, \dots, 1]^T \in \mathbb{R}^n$ will be derived. In the zero-sum stochastic game of agent i , the controller and the adversary are the two players. The strategy space of the control player is $\mathcal{U}_i = \mathbb{R}^n$, while that of the adversarial player (the sensor attack) is explicitly given by

$$\mathcal{Y}_i = \left\{ \Upsilon_i = [\beta_i, 1, \dots, 1]^T \mid \beta_i(x_{i,1}) = \varphi_i(x_{i,1})p_i + k_i, p_i \in \mathbb{R}, k_i \sim \mathcal{N}(0, \sigma_i^2) \right\} \quad (3.38)$$

whose physical meaning follows from the deterministic attack model (2.2). The *worst-case* severe attack is then well defined as the maximizer of the payoff within the admissible set $\mathcal{Y}_i$; it is exactly the saddle-point (maximizing) policy of the zero-sum stochastic game, i.e., $\Upsilon_i^* = \frac{1}{2}A_{ii}^{-1}K_i^T(X_i)\nabla J_i^*(E_i) \in \mathcal{Y}_i$ of (3.42). Letting $E_i = [e_{i,1}, \dots, e_{i,n}]^T \in \mathbb{R}^n$, according to the analysis in the previous section, for the i th agent, construct the consensus error dynamics in the Itô form

$$dE_i = \left(H_i(E_i) + M_i(X_i)U_i + K_i(X_i)\Upsilon_i \right) dt + G_i d\varpi_i \quad (3.39)$$

where

$$\begin{aligned} H_i(E_i) &= [h_{i,1}(e_{i,1}), \dots, h_{i,n}(e_{i,n})]^T \\ M_i(X_i) &= \text{diag}[m_{i,1}, \dots, m_{i,n}] \\ K_i(X_i) &= \text{diag}[(d_i + b_i), 0, \dots, 0] \\ G_i(X_i) &= \text{diag}[g_{i,1}(\bar{x}_{i,1}), \dots, g_{i,n}(\bar{x}_{i,n})]^T \end{aligned}$$

with $M_i(X_i), K_i(X_i) \in \mathbb{R}^{n \times n}$, $G_i(X_i) \in \mathbb{R}^{n \times r_i}$, and $H_i(E_i) \in \mathbb{R}^n$.

Definition 2. Consider the communication graph $\bar{\mathcal{G}}_r$. The graphical game is defined as

$$\mathcal{G}_g = (\mathcal{P}, \{\mathcal{U}_i\}_{i=1}^N, \{J_i\}_{i=1}^N),$$

where $\mathcal{P} = \{1, 2, \dots, N\}$ denotes the set of follower agents. For each player $i \in \mathcal{P}$, the distributed feedback strategy is given by $U_i(E_i) \in \mathcal{U}_i = \mathbb{R}^n$, while the associated adversary selects the attack strategy $\Upsilon_i \in \mathcal{Y}_i$ defined in (3.38).

The information available to player i consists of their own consensus error E_i and the local neighboring information E_j , $j \in \mathcal{N}_i$, through the communication graph. Therefore, no global information about the entire MAS is required. The payoff of player i is described by the minimax performance functional $J_i(E_i)$ in (3.40), where the graphical coupling among neighboring players is characterized by $\sum_{j \in \mathcal{N}_i} U_j^T R_{ij} U_j$ and $\sum_{j \in \mathcal{N}_i} \Upsilon_j^T A_{ij} \Upsilon_j$. A strategy profile $\{U_i^*\}_{i=1}^N$ is called a Nash equilibrium of the graphical game if, for every $i \in \mathcal{P}$,

$$J_i(U_i^*, U_{-i}^*) \leq J_i(U_i, U_{-i}^*), \quad \forall U_i \in \mathcal{U}_i.$$

Thus, no player can improve its own payoff by unilaterally deviating from its equilibrium strategy. The local minimax solution (3.42) of the zero-sum stochastic game, coupled through the graphical payoff, yields the network-level Nash equilibrium.

According to Definition 2, the local zero-sum games are coupled through the communication topology and collectively constitute the network-level graphical game.

Next, a common assumption is introduced that is used in numerous related studies [34].

Assumption 3. [34] The Nash equilibrium solution of a stochastic game is bounded $H_i(E_i) + M_i(X_i)U_i^* + K_i(X_i)\Upsilon_i^* \leq \iota_i \|E_i\|$ with $\iota_i > 0$.

Remark 3. Given the graphical-game coupling, the individual HJI equations (3.47), $i = 1, \dots, N$, form a set of coupled HJI equations. Under Assumption 3, which postulates the existence of a bounded and admissible Nash solution, and within the standard continuous-time policy-iteration framework [34], the coupled HJIs admit a jointly bounded, admissible solution $\{J_i^*\}_{i=1}^N$ on the compact set. Because each agent minimizes only its own local utility function that is coupled with its neighbors through the terms $\sum_{j \in \mathcal{N}_i} U_j^T R_{ij} U_j$ and $\sum_{j \in \mathcal{N}_i} \Upsilon_j^T A_{ij} \Upsilon_j$ in (3.40), the local (agent-level) equilibria that are synchronized through the graphical-game coupling constitute a fixed point of the coupled HJI system, which certifies the global Nash equilibrium.

The control goal of this paper is to seek the Nash equilibrium solution (U_i^*, Υ_i^*) while minimizing the control input U_i^* and maximizing the sensor attack Υ_i^* . To accomplish this goal, the minimax optimal control method is employed, particularly under worst-case attack scenarios, ensuring the attainment of the optimal performance function as follows:

$$\begin{aligned} J_i(E_i) &= \min_{U_i} \max_{\Upsilon_i} \left[\mathbf{E} \left\{ \int_0^\infty r_i(E_i, U_i, \Upsilon_i) d\tau \right\} \right] \\ &= \min_{U_i} \max_{\Upsilon_i} \left[\mathbf{E} \left\{ \int_0^\infty \left(Q_i(E_i(\tau)) + U_i^T R_{ii} U_i - \Upsilon_i^T A_{ii} \Upsilon_i + \sum_{j \in \mathcal{N}_i} U_j^T R_{ij} U_j - \sum_{j \in \mathcal{N}_i} \Upsilon_j^T A_{ij} \Upsilon_j \right) d\tau \right\} \right] \end{aligned} \quad (3.40)$$

where $\mathbf{E}\{\cdot\}$ denotes the expectation. $Q_i(E_i)$ denotes the semi-definite penalty function. $R_{ii} \in \mathbb{R}^{n \times n} > 0$, $A_{ii} \in \mathbb{R}^{n \times n} > 0$, $R_{ij} \in \mathbb{R}^{n \times n} \geq 0$, and $A_{ij} \in \mathbb{R}^{n \times n} \geq 0$ represent positive symmetric matrices.

Subsequently, the Hamiltonian function of the related cost function (3.40) is defined as

$$H_a(E_i, U_i, \Upsilon_i)$$

$$\begin{aligned}
&= (\nabla J_i(E_i))^T \left(H_i(E_i) + M_i(X_i)U_i + K_i(X_i)\Upsilon_i \right) + \frac{1}{2} \text{Tr} \left\{ G_i^T \nabla^2 J_i(E_i) G_i \right\} \\
&\quad + Q_i(E_i(\tau)) + U_i^T R_{ii} U_i - \Upsilon_i^T A_{ii} \Upsilon_i + \sum_{j \in N_i} U_j^T R_{ij} U_j - \sum_{j \in N_i} \Upsilon_j^T A_{ij} \Upsilon_j
\end{aligned} \tag{3.41}$$

where $\nabla J_i(E_i)$ denotes the gradient of $J_i(E_i)$.

Then, the optimal control scheme (U_i^*, Υ_i^*) is formulated as

$$\begin{cases} U_i^* = -\frac{1}{2} R_{ii}^{-1} M_i^T(X_i) \nabla J_i^*(E_i) \\ \Upsilon_i^* = \frac{1}{2} A_{ii}^{-1} K_i^T(X_i) \nabla J_i^*(E_i). \end{cases} \tag{3.42}$$

where the definition of $\nabla J_i^*(E_i)$ is given below.

According to the error dynamics (3.39), cost function (3.40), and the optimal scheme (3.42), a Lyapunov candidate function $J_{si}(E_i)$ is designed, satisfying

$$\mathcal{L}J_{si}(E_i) = (\nabla J_{si}(E_i))^T (H_i(E_i) + M_i(X_i)U_i^* + K_i(X_i)\Upsilon_i^*) + \frac{1}{2} \text{Tr} \left\{ G_i^T \nabla^2 J_{si}(E_i) G_i \right\} < 0.$$

where $\nabla J_{si}(E_i)$ is the gradient of $J_{si}(E_i)$. Furthermore, define a positive definite function $\Lambda_i(E_i)$ where $\forall E_i \neq 0, \Lambda_i(E_i) > 0$, such that $E_i = 0 \implies \Lambda_i(E_i) = 0$. For $\Lambda_i(E_i)$, there is $\lim_{E_i \rightarrow \infty} \Lambda_i(E_i) = \infty$ and $(\nabla J_i^*(E_i))^T \Lambda_i(E_i) \nabla J_{si}(E_i) = r_i(E_i, U_i^*, \Upsilon_i^*)$. Then, the following relation holds:

$$\mathcal{L}J_{si}(E_i) \leq -(\nabla J_{si}(E_i))^T \Lambda_i(E_i) \nabla J_{si}(E_i). \tag{3.43}$$

Subsequently, to achieve the optimal strategy for a stochastic game (3.42), FLSs are used to approximate the consensus cost function

$$J_i^*(E_i) = W_{ci}^T \delta_i(E_i) + \epsilon_{ci} \tag{3.44}$$

where $W_{ci} \in \mathbb{R}^{L_0}$, $\delta_i(E_i) \in \mathbb{R}^{L_0}$, and $\epsilon_{ci} \in \mathbb{R}$ are the ideal weight, the basis function vector, and the approximate error of FLSs, respectively.

The gradient of (3.44) is calculated as

$$\nabla J_i^*(E_i) = (\nabla \delta_i(E_i))^T W_{ci} + \nabla \epsilon_{ci}. \tag{3.45}$$

Therefore, (3.42) can be rewritten as

$$\begin{cases} U_i^* = -\frac{1}{2} R_{ii}^{-1} M_i^T(X_i) \left((\nabla \delta_i(E_i))^T W_{ci} + \nabla \epsilon_{ci} \right) \\ \Upsilon_i^* = \frac{1}{2} A_{ii}^{-1} K_i^T(X_i) \left((\nabla \delta_i(E_i))^T W_{ci} + \nabla \epsilon_{ci} \right). \end{cases} \tag{3.46}$$

Since $J_i^*(E_i)$ is independent of the realization of the Gaussian noise and the diffusion matrix G_i is independent of U_i and Υ_i , the second-order diffusion term in the Hamiltonian is a state-dependent constant with respect to the players' actions. Consequently, the first-order stationarity conditions $\frac{\partial H_a}{\partial U_i} =$

0 and $\frac{\partial H_a}{\partial \Upsilon_i} = 0$ are not affected by the diffusion term, which explains why the noise intensity $g_{i,s}$ does not appear explicitly in the optimal law (3.42) but only affects the value of the performance function.

Let $\mathcal{D}_i(E_i) = \frac{1}{2} \left[\text{Tr}\{G_i^T \nabla^2 \delta_{i,1}(E_i) G_i\}, \dots, \text{Tr}\{G_i^T \nabla^2 \delta_{i,L_0}(E_i) G_i\} \right]^T$, and $d_{\varepsilon i}(E_i) = \frac{1}{2} \text{Tr}\{G_i^T \nabla^2 \varepsilon_{ci}(E_i) G_i\}$. The HJI equation is obtained as

$$H_a(E_i, W_{ci}, U_i^*, \Upsilon_i^*) = Q_i(E_i) + W_{ci}^T [\nabla \delta_i(E_i) H_i(E_i) + \mathcal{D}_i(E_i)] - \frac{1}{4} W_{ci}^T G_{1i} W_{ci} + \frac{1}{4} \sum_{j \in \mathcal{N}_i} W_{cj}^T G_{2j} W_{cj} + \delta_{HJI} = 0. \quad (3.47)$$

where

$$G_{1i} = \nabla \delta_i(E_i) (\bar{M}_i - \bar{K}_i) \nabla \delta_i(E_i)^T \\ G_{2j} = \nabla \delta_j(E_j) (\bar{M}_{-i} - \bar{K}_{-i}) \nabla \delta_j(E_j)^T$$

with

$$\bar{M}_i = M_i(X_i) R_{ii}^{-1} M_i^T(X_i) \\ \bar{K}_i = K_i(X_i) A_{ii}^{-1} K_i^T(X_i) \\ \bar{M}_{-i} = M_j R_{jj}^{-1} R_{ij} R_{jj}^{-1} M_j^T(X_j) \\ \bar{K}_{-i} = K_j(X_j) A_{jj}^{-1} A_{ij} A_{jj}^{-1} K_j^T(X_j).$$

$\bar{M}_i, \bar{K}_i, \bar{M}_{-i}, \bar{K}_{-i} \in \mathbb{R}^{n \times n}$, and the matrices $G_{1i}, G_{2j} \in \mathbb{R}^{L_0 \times L_0}$. The residual error δ_{HJI} is given by

$$\delta_{HJI} = (\nabla \varepsilon_{ci})^T \left(H_i(E_i) + M_i(X_i) U_i^* + K_i(X_i) \Upsilon_i^* \right) + \frac{1}{4} (\nabla \varepsilon_{ci})^T (\bar{M}_i - \bar{K}_i) \nabla \varepsilon_{ci} \\ + \sum_{j \in \mathcal{N}_i} \left[\frac{1}{2} W_{cj}^T \nabla \delta_j(E_j) (\bar{M}_{-i} - \bar{K}_{-i}) \nabla \varepsilon_{cj} + \frac{1}{4} (\nabla \varepsilon_{cj})^T (\bar{M}_{-i} - \bar{K}_{-i}) \nabla \varepsilon_{cj} \right] \\ + d_{\varepsilon i}(E_i). \quad (3.48)$$

However, the ideal W_{ci} is uncertain. Therefore, the consensus cost function is approximated as

$$\hat{J}_i(E_i) = \hat{W}_{ci}^T \delta_i(E_i) \quad (3.49)$$

where $\hat{W}_{ci}$ denotes the estimation of W_{ci} , and $\delta_i(E_i)$ is the same basis-function vector as in (3.44). Then, the optimal scheme (3.42) is estimated as

$$\begin{cases} \hat{U}_i = -\frac{1}{2} R_{ii}^{-1} M_i^T(X_i) (\nabla \delta_i(E_i))^T \hat{W}_{ci} \\ \hat{\Upsilon}_i = \frac{1}{2} A_{ii}^{-1} K_i^T(X_i) (\nabla \delta_i(E_i))^T \hat{W}_{ci}. \end{cases} \quad (3.50)$$

Similarly, design an approximate Hamiltonian function

$$H_a(E_i, \hat{W}_{ci}, \hat{U}_i, \hat{\Upsilon}_i) = Q_i(E_i) + \hat{W}_{ci}^T [\nabla \delta_i(E_i) H_i(E_i) + \mathcal{D}_i(E_i)]$$

$$-\frac{1}{4}\hat{W}_{ci}^T G_{1i} \hat{W}_{ci} + \frac{1}{4} \sum_{j \in N_i} \hat{W}_{cj}^T G_{2j} \hat{W}_{cj} \triangleq e_{ci} \quad (3.51)$$

Obviously, the approximate Hamiltonian function $H_a(E_i, \hat{W}_{ci}, \hat{U}_i, \hat{Y}_i)$ can be minimized by adjusting the estimated weight $\hat{W}_{ci}$, i.e., the updating law designed $\dot{\hat{W}}_{ci}$ can minimize the objective function $E_{ci} = \frac{1}{2} e_{ci}^T e_{ci}$.

Using the gradient descent method, the update law of the FLS weight is constructed as

$$\begin{aligned} \dot{\hat{W}}_{ci} = & -\frac{\alpha_i \phi_i}{m_{si}^2} e_{ci} + \frac{\alpha_i}{2} \nabla \delta_i(E_i) (\bar{M}_i - \bar{K}_i) \nabla J_{si}(E_i) \\ & + \alpha_i \left[\frac{\phi_i}{4m_{si}^2} \hat{W}_{ci}^T G_{1i} \hat{W}_{ci} - \sum_{j \in N_i} \frac{\phi_i}{4m_{si}^2} \hat{W}_{cj}^T G_{2j} \hat{W}_{cj} - (F_{2i} \hat{W}_{ci} - F_{1i} \varphi_i^T \hat{W}_{ci}) \right] \end{aligned} \quad (3.52)$$

where $\alpha_i > 0$ is a learning parameter, $\phi_i = \nabla \delta_i(E_i) (H_i(E_i) + M_i(X_i) \hat{U}_i) + \mathcal{D}_i(E_i) \in \mathbb{R}^{L_0}$, $m_{si} = 1 + \phi_i^T \phi_i \in \mathbb{R}$ is introduced for normalization with $\varphi_i = \phi_i / m_{si}$, and $F_{1i} \in \mathbb{R}^{L_0}$, $F_{2i} \in \mathbb{R}^{L_0 \times L_0}$ are suitable tuning parameters.

In order to improve the security performance of stochastic MASs over sensor attacks, an optimal CSC algorithm based on the stochastic game is designed. By designing the critic weight update laws (3.52) based on FLSs, the policy iteration will be more accurate.

Remark 4. Although the existing results [13, 15, 20, 35] have considered deception attacks and guaranteed the stability of closed-loop systems, the limitations of these works are that the stability is guaranteed without achieving optimality, and the worst-case attack case is not taken into account. Therefore, this paper embeds the stochastic game algorithm into the backstepping technique to develop an optimal CSC strategy that minimizes the cost function and ensures the stability of the closed-loop system even in the worst attack scenarios.

4. Stability analysis

Theorem 1. Consider uncertain stochastic MASs with sensor attacks over unreliable communication networks (2.1), and design the distributed controllers (3.16), (3.26), and (3.33) with adaptive laws (3.17), (3.27), and (3.34). According to the proposed stochastic game-based CSC protocol, based on the cost function (3.40), and the weight update law (3.52). Selecting the design parameters ensures that consensus errors $e_{i,s}$, estimation errors $\tilde{\theta}_{i,s}$, and the approximation error $\tilde{W}_{ci}$ of the updated law are mean-square uniformly ultimately bounded (MS-UUB).

Proof. For closed-loop systems, construct the Lyapunov function as

$$V_{HJI} = V_S + \sum_{i=1}^N \left[\frac{1}{2} \tilde{W}_{ci}^T \alpha_i^{-1} \tilde{W}_{ci} + J_{si}(E_i) \right]. \quad (4.1)$$

The derivative of V_{HJI} is calculated as

$$\mathcal{L}V_{HJI} = \mathcal{L}V_S + \sum_{i=1}^N \left[\tilde{W}_{ci}^T \alpha_i^{-1} \dot{\tilde{W}}_{ci} + \mathcal{L}J_{si}(E_i) \right]. \quad (4.2)$$

Considering (3.37), let $E_i = [e_{i,1}, \dots, e_{i,n}]^T$ and $\tilde{\theta}_i = [\tilde{\theta}_{i,1}, \dots, \tilde{\theta}_{i,n}]^T$. Then we have

$$\begin{aligned} \mathcal{L}V_{HJI} = & \sum_{i=1}^N \left[E_i^3 (H_i(E_i) + M_i(X_i)U_i^* + K_i(X_i)\Upsilon_i^*) \right] \\ & - \sum_{i=1}^N \left(\sum_{p=1}^{n-1} \bar{\tau}_{i,p+1} z_{i,p+1}^4 + c_i \|E_i\|^4 + \bar{\gamma}_i \|\tilde{\theta}_i\|^2 \right) \\ & + \sum_{i=1}^N \left[\tilde{W}_{ci}^T \alpha_i^{-1} \dot{\tilde{W}}_{ci} + \mathcal{L}J_{si}(E_i) + P_i \right]. \end{aligned} \quad (4.3)$$

where

$$\begin{aligned} \bar{\tau}_{i,p+1} &= \frac{1}{\tau_{i,p}} - \frac{m_{i,p}^4}{4} - \frac{3}{4} (\kappa_{i,p} B_{i,p+1})^{\frac{4}{3}} \\ c_i &= \min\{c_{i,p} | 1 \leq p \leq n\} \\ \bar{\gamma}_i &= \min\{\bar{\gamma}_{i,p} | 1 \leq p \leq n\} \text{ with } \bar{\gamma}_{i,p} = \frac{\gamma_{i,p}}{2\mathcal{Q}_{i,p}}. \end{aligned}$$

Let $\tilde{W}_{ci} = W_{ci} - \hat{W}_{ci}$ and $Y_i = [\varphi_i^T \tilde{W}_{ci}, \tilde{W}_{ci}]^T$. According to (3.47), (3.51) and (3.52), the error dynamics are calculated as

$$\tilde{W}_{ci}^T \alpha_i^{-1} \dot{\tilde{W}}_{ci} = -Y_i^T \Psi_i Y_i + Y_i^T \Xi_i - \frac{1}{2} \tilde{W}_{ci}^T \nabla \delta_i(E_i) (\bar{M}_i - \bar{K}_i) \nabla J_{si}(E_i) \quad (4.4)$$

where only the symmetric part of Ψ_i contributes to the quadratic form $Y_i^T \Psi_i Y_i$. Therefore, define

$$\Psi_i^s = \frac{1}{2} (\Psi_i + \Psi_i^T) = \begin{bmatrix} I & -\frac{1}{2} F_{1i}^T \\ -\frac{1}{2} F_{1i} & B_i^s \end{bmatrix},$$

where

$$B_i^s = \frac{1}{2} (F_{2i} + F_{2i}^T) - \frac{1}{4} \left[\frac{\phi_i}{m_{si}^2} W_{ci}^T G_{1i} + G_{1i}^T W_{ci} \frac{\phi_i^T}{m_{si}^2} \right].$$

Thus, $Y_i^T \Psi_i Y_i = Y_i^T \Psi_i^s Y_i$.

Lemma 2. Suppose that $F_{2i} = F_{2i}^T$ and that, on the compact set under consideration,

$$\|W_{ci}\| \leq W_{Mi}, \quad \|G_{1i}\| \leq \bar{G}_{1i}, \quad \frac{\|\phi_i\|}{m_{si}^2} \leq \bar{\rho}_i.$$

If F_{2i} is selected such that

$$\lambda_{\min}(F_{2i}) > \frac{1}{4} \|F_{1i}\|^2 + \frac{1}{2} \bar{\rho}_i W_{Mi} \bar{G}_{1i},$$

then Ψ_i^s is positive definite.

Proof. By the Schur complement, $\Psi_i^s > 0$ is guaranteed if

$$B_i^s - \frac{1}{4} F_{1i} F_{1i}^T > 0.$$

Moreover,

$$\left\| \frac{1}{4} \left[\frac{\phi_i}{m_{si}^2} W_{ci}^T G_{1i} + G_{1i}^T W_{ci} \frac{\phi_i}{m_{si}^2} \right] \right\| \leq \frac{1}{2} \bar{\rho}_i W_{Mi} \bar{G}_{1i}.$$

Therefore, the above condition on F_{2i} guarantees $\Psi_i^s > 0$, and consequently

$$\lambda_{\min}(\Psi_i^s) > 0.$$

By Lemma 2, the parameters F_{1i} and F_{2i} are selected to ensure the positive definiteness of the symmetric matrix Ψ_i^s , and (4.3) is further calculated as

$$\begin{aligned} \mathcal{L}V_{HJI} \leq & \sum_{i=1}^N \left[\mathcal{L}J_{si}(E_i) - \sum_{p=1}^{n-1} \bar{\tau}_{i,p+1} z_{i,p+1}^4 - c_i \|E_i\|^4 \right] \\ & + \sum_{i=1}^N \left[-\lambda_{\min}(\Psi_i^s) \|Y_i\|^2 + \|\Xi_i\| \|Y_i\| - \bar{\gamma}_i \|\tilde{\theta}_i\|^2 \right] \\ & + \sum_{i=1}^N \left[E_i^3 (H_i(E_i) + M_i(X_i)U_i^* + K_i(X_i)\Upsilon_i^*) \right] \\ & + \sum_{i=1}^N \left[P_i - \frac{1}{2} \tilde{W}_{ci}^T \nabla \delta_i(E_i) (\bar{M}_i - \bar{K}_i) \nabla J_{si}(E_i) \right] \end{aligned} \quad (4.5)$$

According to Assumption 3 and Young's inequality, there is

$$E_i^3 (H_i(E_i) + M_i(X_i)U_i^* + K_i(X_i)\Upsilon_i^*) \leq \frac{3}{4\varsigma_i^2} \|E_i\|^4 + \frac{\varsigma_i^2 \iota_i^2}{4} \|E_i\|^2 \quad (4.6)$$

where $\varsigma_i > 0$ is a design parameter.

Since M_i and K_i are bounded, $M_i - K_i$ is also bounded, and $\|M_i - K_i\| \leq \kappa_i$, where κ_i is a positive constant. Let $z_i = [z_{i,2}, \dots, z_{i,n}]^T$ and $\bar{\tau}_i = \min\{\bar{\tau}_{i,p+1} | 1 \leq p \leq n-1\}$. Based on Assumption 3 and (3.43), (4.5) yields

$$\begin{aligned} \mathcal{L}V_{HJI} \leq & \sum_{i=1}^N \left[-\bar{c}_i \|E_i\|^4 + \frac{\varsigma_i^2 \iota_i^2}{4} \|E_i\|^2 + \|\Xi_i\| \|Y_i\| - \bar{\tau}_i \|z_i\|^4 - \lambda_{\min}(\Psi_i^s) \|Y_i\|^2 \right. \\ & \left. + \frac{1}{2} \kappa_i \epsilon_{Mi} \|\nabla J_{si}(E_i)\| + P_i^s - \lambda_{\min} \Lambda_i(E) \|\nabla J_{si}(E_i)\|^2 - \bar{\gamma}_i \|\tilde{\theta}_i\|^2 \right] \end{aligned} \quad (4.7)$$

where $\bar{c}_i = c_i - \frac{3}{4\varsigma_i^2}$ and the bounded Itô second-order term $\frac{1}{2} \text{Tr}\{G_i^T \nabla^2 J_{si}(E_i) G_i\} \leq \bar{P}_i$ has been absorbed into the unified constant $P_i^s = P_i + \bar{P}_i$

To achieve $\mathcal{L}V_{HJI} < 0$, the parameter in (4.7) must satisfy the following conditions:

$$\|E_i\|^2 \geq \sqrt{\frac{\varsigma_i^4 \iota_i^4}{64\bar{c}_i^2} + \frac{P_i^s}{\bar{c}_i}} + \frac{\varsigma_i^2 \iota_i^2}{8\bar{c}_i}$$

or

$$\|\tilde{\theta}_i\| \geq \sqrt{\frac{P_i^s}{\tilde{\gamma}_i}}$$

or

$$\|z_i\|^2 \geq \sqrt{\frac{P_i^s}{\tilde{\tau}_i}}$$

or

$$\|Y_i\| \geq \sqrt{\frac{\|\Xi_i\|^2}{4\lambda_{\min}^2(\Psi_i^s)} + \frac{P_i^s}{\lambda_{\min}(\Psi_i^s)}} + \frac{\|\Xi_i\|}{2\lambda_{\min}(\Psi_i^s)}$$

or

$$\|\nabla J_{si}(E_i)\| \geq \sqrt{\frac{\kappa_i^2 \epsilon_{\tilde{M}i}^2}{16\lambda_{\min}^2(\Lambda_i(E_i))} + \frac{P_i^s}{\lambda_{\min}(\Lambda_i(E_i))}} + \frac{\kappa_i \epsilon_{\tilde{M}i}}{4\lambda_{\min}(\Lambda_i(E_i))}.$$

In particular, since $Y_i = [\varphi_i^T \tilde{W}_{ci}, \tilde{W}_{ci}]^T$, the bound on $\|Y_i\|$ in the above conditions directly yields an explicit ultimate bound on the weight-approximation error,

$$\|\tilde{W}_{ci}\| \leq \bar{W}_{ci} := \sqrt{\frac{\|\Xi_i\|^2}{4\lambda_{\min}^2(\Psi_i^s)} + \frac{P_i^s}{\lambda_{\min}(\Psi_i^s)}} + \frac{\|\Xi_i\|}{2\lambda_{\min}(\Psi_i^s)}.$$

Since the above bounds are established through the Itô derivative of the Lyapunov function V_{HII} and the expectation $\mathbf{E}\{V_{HII}\}$, the resulting stability is MS-UUB, i.e.,

$$\mathbf{E}[\|E_i(t)\|^2 + \|\tilde{\theta}_i(t)\|^2 + \|\tilde{W}_{ci}(t)\|^2] \leq \Delta, \quad \forall t \geq T,$$

for some finite T and constant Δ .

By applying the Lyapunov theorem, it is demonstrated that all error signals within the closed-loop systems are MS-UUB. The proof is completed. $\square$

Remark 5. The Nash equilibrium is an optimality condition described by the coupled HJI equations (3.47) and the minimax policy (3.42). The stability of the closed loop guarantees that the proposed control protocol is physically realizable, while the graphical-game equilibrium, whenever it is jointly solvable (see Remark 3), ensures the network-level optimality in the sense of expectation.

5. Simulation studies

The proposed stochastic game-based CSC scheme will verify the validity through two simulation examples.

5.1. Numerical example

A set of stochastic MASs is considered to consist of a leader and four followers, according to the topology shown in Figure 2. The dynamics of the followers are as follows:

$$dx_{i,1} = (0.9x_{i,2} + 8 \sin(x_{i,1})) dt + 0.1 \sin(x_{i,1}) d\omega$$

$$dx_{i,2} = (0.8u_i + 8 \sin(x_{i,1} - 0.3)) dt + 0.1 \cos(x_{i,1}) d\omega$$

$$y_i = x_{i,1} + \beta_i(x_{i,1})$$

where $\beta_i(x_{i,1}) = \varphi_i(x_{i,1}) + k_{bias}$, $\varphi_i(x_{i,1}) = -0.45x_{i,1}$ represents the regressor function, and k_{bias} denotes the attack bias, which is composed of $\delta_{imp}(0.1\sin(t))$, with δ_{imp} denoting a pulse signal function. The design parameters of the control protocol satisfy $c_{i,1} = 9$, $c_{i,2} = 9$, $a_{i,1} = 15$, $a_{i,2} = 20$, $\varrho_{i,1} = 30$, $\varrho_{i,2} = 90$, $\gamma_{i,1} = 50$, $\gamma_{i,2} = 150$, and $\tau_{i,1} = 0.01$.

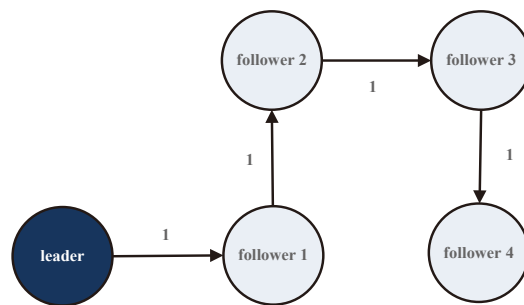


Figure 2. Communication graph.

Case 1: The simulation for the stochastic game-based optimal CSC scheme.

The relevant initial values of the stochastic game strategy are selected as 0. The matrices are designed as follows: $R_{11} = R_{22} = R_{33} = R_{44} = 10I$, $R_{21} = R_{32} = R_{43} = I$, $A_{11} = A_{22} = A_{33} = A_{44} = 10I$, $A_{21} = A_{32} = A_{43} = I$. Select the tuning parameters $F_{11} = F_{12} = F_{13} = F_{14} = 120 \cdot [1, 1, 1, 1]^T$, $F_{21} = F_{22} = F_{23} = F_{24} = 200I$. The learning law is $\alpha_i = 0.4$. The Lyapunov function $J_{si}(E_i)$ is constructed as $J_{si}(E_i) = \frac{1}{2} (e_{i,1}^2 + e_{i,2}^2)$, $i = 1, 2, 3, 4$.

According to the designed optimal CSC strategy, the consensus trajectories of the four follower agents and the leader agent are shown in Figure 3. It is evident that all the followers successfully achieve consensus tracking. Figure 4 displays the consensus tracking errors $e_{i,1}$, which are ensured to remain bounded within a compact set. Figure 5 presents the state trajectories of four agents x_{i2} . Furthermore, Figure 6 displays the CSC action against sensor attacks.

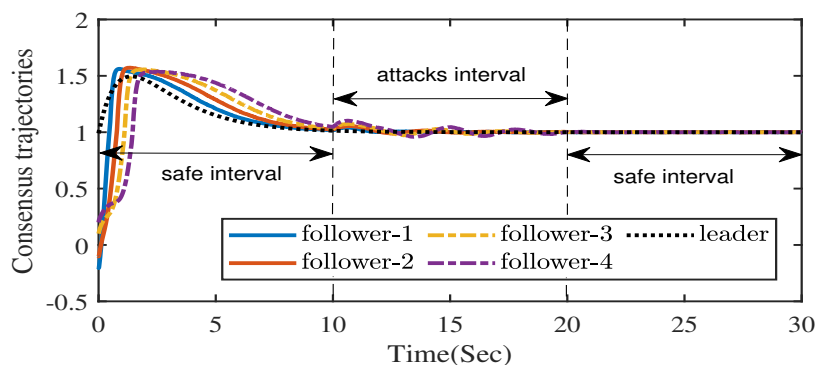


Figure 3. The leader-following consensus trajectories in Case 1.

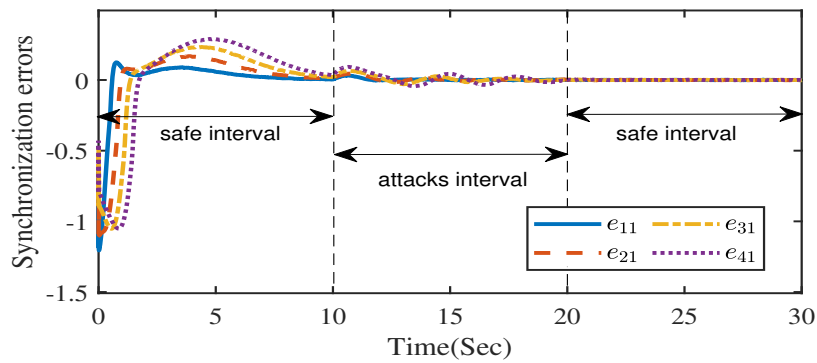


Figure 4. Synchronization errors of the four agents in Case 1.

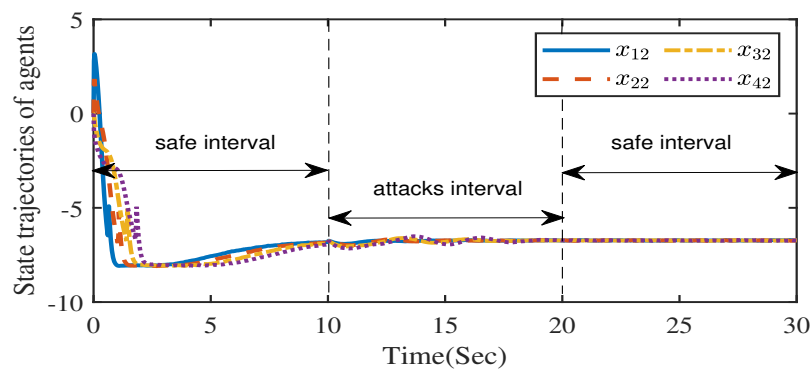


Figure 5. State trajectories of the four agents in Case 1.

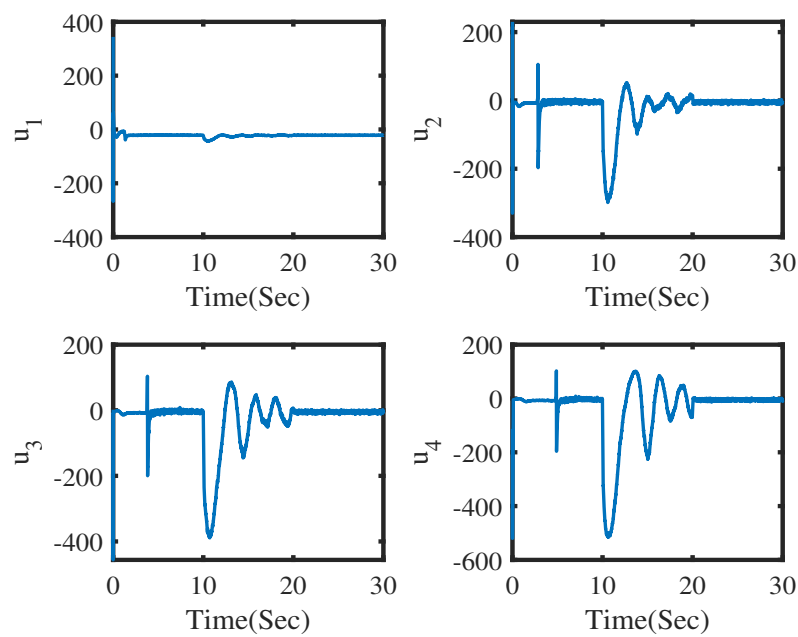


Figure 6. Control inputs of the four agents in Case 1.

Case 2: The comparison experiment.

In this section, the existing adaptive security control method [20–22] is adopted, that is, the stochastic game strategy is not used to compare the control performance of the system under sensor attacks, and other parameters remain unchanged except the comparison variables. Since we are mainly concerned with the consensus tracking control performance of the system, only the consensus trajectories and synchronization errors are given in this section, as shown in Figures 7 and 8, respectively. Compared with Figure 7, it can be concluded that Figure 3 in Case 1 has better consensus control effect during the sensor attack interval. Similarly, compared with Figure 8, it is evident that Figure 4 in Case 1 exhibits smaller synchronization errors during the sensor attack period. This further validates that the proposed optimal CSC scheme based on stochastic game theory yields superior control performance for stochastic MASs over sensor attacks. Figure 9 shows the weights $\hat{W}_{ci}$ ($i = 1, 2, 3, 4$) of the four intelligent agents.

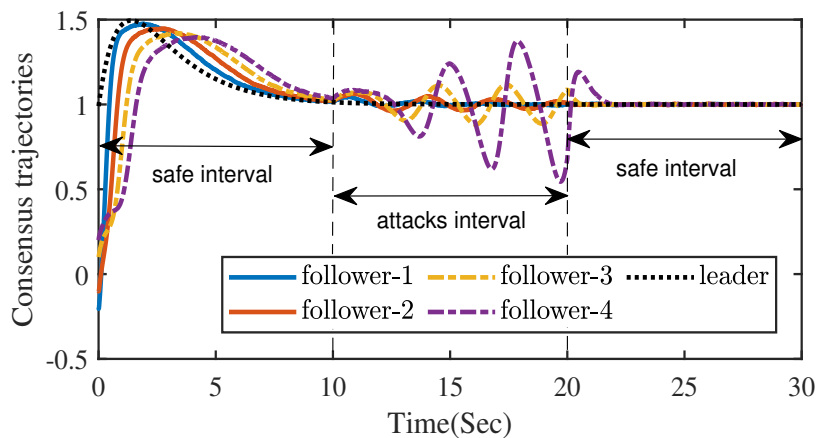


Figure 7. The leader-following consensus trajectories in Case 2.

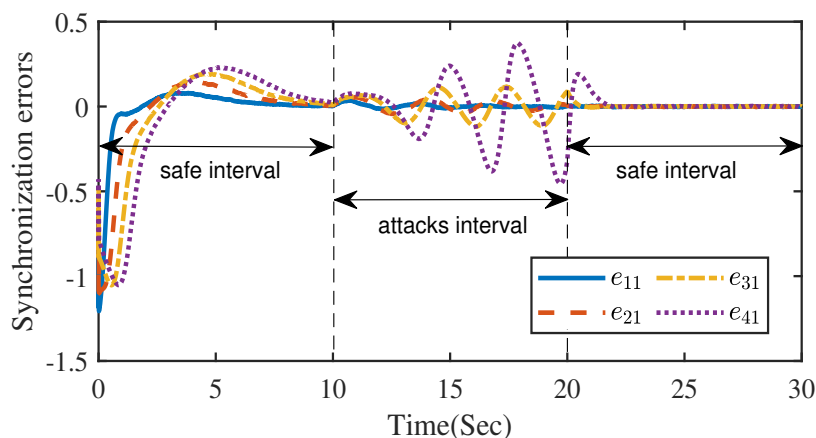


Figure 8. Synchronization errors of the four agents in Case 2.

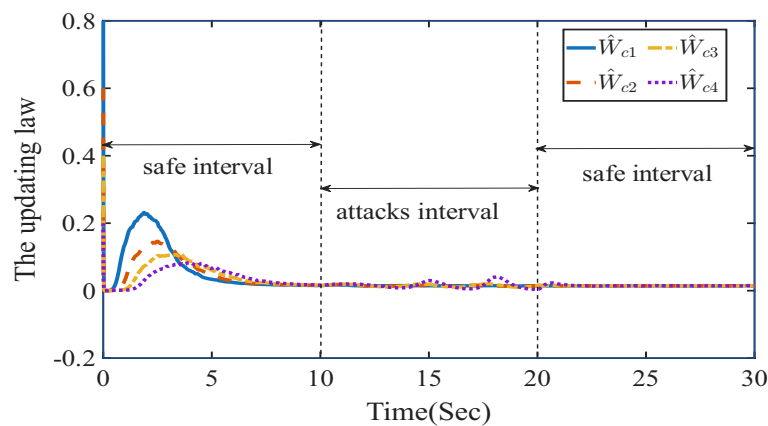


Figure 9. The updating laws of the four agents.

5.2. Practical example

To evaluate the effectiveness of the control method, the islanded microgrids test system [36] is considered. The islanded microgrid test consists of four inverter-based distributed generations (DGs) units. Figure 10 indicates the considered communication topology and one-line diagram of the islanded microgrids test system, where only DG 1 can receive the reference signal.

For each inverter-based DG, the voltage/current dynamics in the dq -rotation frame is first transformed into the strict-feedback stochastic form (2.1) with the following physical mapping: the d -axis voltage-tracking error of DG i is taken as $x_{i,1}$, the auxiliary voltage variable as $x_{i,2}$, the inverter control input as u_i , and the measured voltage with the sensor attack as $y_i = x_{i,1} + \beta_i(x_{i,1})$. The corresponding coefficients are $m_{i,1}$ (droop/inertia gain), $m_{i,2}$ (local DG gain), $q_{i,1}(\bar{x}_{i,1})$ (droop and load-current nonlinearity), $q_{i,2}(\bar{x}_{i,2})$ (inner voltage/current loop nonlinearity), and $g_{i,1}, g_{i,2}$ (load/noise coefficients). The grid and load parameters, the communication topology and the reference value are given below so that the example is reproducible as an engineering validation.

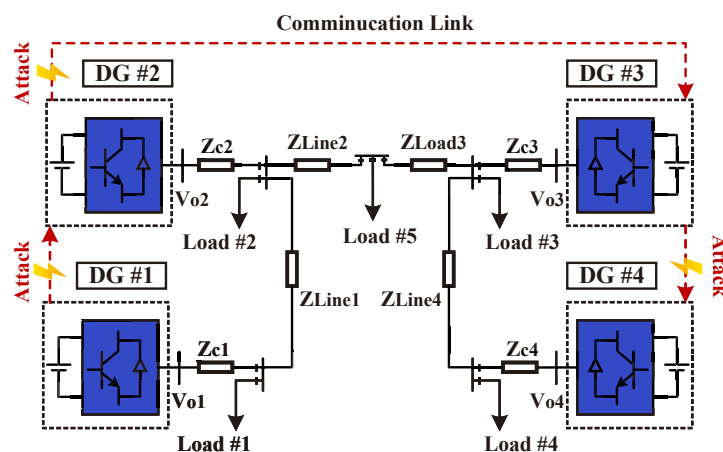


Figure 10. Islanded microgrids test system.

Select the voltage reference value $y_r = 10\cos(t)(-\exp(-t)) + 10$. The sensor attack model is the same as the numerical example. The design parameters of the control protocol satisfy $c_{i,1} = 4$, $c_{i,2} = 60$, $a_{i,1} = 15$, $a_{i,2} = 20$, $q_{i,1} = 30$, $q_{i,2} = 90$, $\gamma_{i,1} = 50$, $\gamma_{i,2} = 150$, and $\tau_{i,1} = 0.01$.

The relevant initial values of the stochastic game strategy are selected as 0. The matrices are designed as $R_{11} = R_{22} = R_{33} = R_{44} = I$, $R_{21} = R_{32} = R_{43} = 0.1I$, $A_{11} = A_{22} = A_{33} = A_{44} = 10I$, $A_{21} = A_{32} = A_{43} = I$. Select the tuning parameters $F_{11} = F_{12} = F_{13} = F_{14} = 200 \cdot [1, 1, 1, 1]^T$, $F_{21} = F_{22} = F_{23} = F_{24} = 20I$. The learning law is $\alpha_i = 0.4$. The Lyapunov function $J_{si}(E_i)$ is constructed as $J_{si}(E_i) = \frac{1}{2}(e_{i,1}^2 + e_{i,2}^2)$, $i = 1, 2, 3, 4$.

Using the designed optimal CSC scheme in island microgrids, the voltage output signals of the four DGs and the voltage reference signal are depicted in Figure 11, where the system is subjected to sensor attacks at 5 seconds and 8 seconds, respectively. It is evident that the islanded microgrids test system under sensor attacks successfully achieves voltage recovery. Figure 12 displays the synchronization errors $e_{i,1}$ between the voltage value of the DG output and the reference value. Figure 13 displays the voltage recovery controllers against sensor attack. By using the proposed distributed stochastic game strategy, Figure 14 presents the evolution of critic network weights $\hat{W}_{ci}$, demonstrating that these weights remain bounded. As the stability of the FLS weights W_{ci} is a prerequisite for applying the FLS to approximate the cost function, Figure 14 indeed shows that the estimated critic weights $\hat{W}_{ci}$ converge to small neighborhoods of constant values and remain bounded, which verifies the applicability of the FLSs and is consistent with the boundedness of $\tilde{W}_{ci}$ established in the proof of Theorem 1.

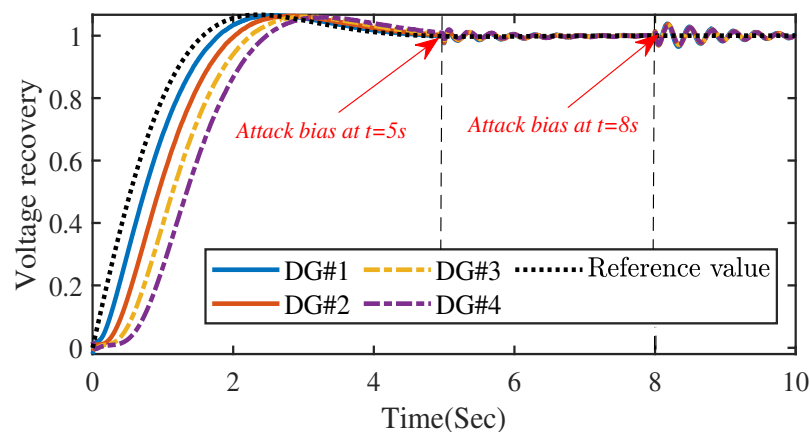


Figure 11. Voltage recovery trajectories.

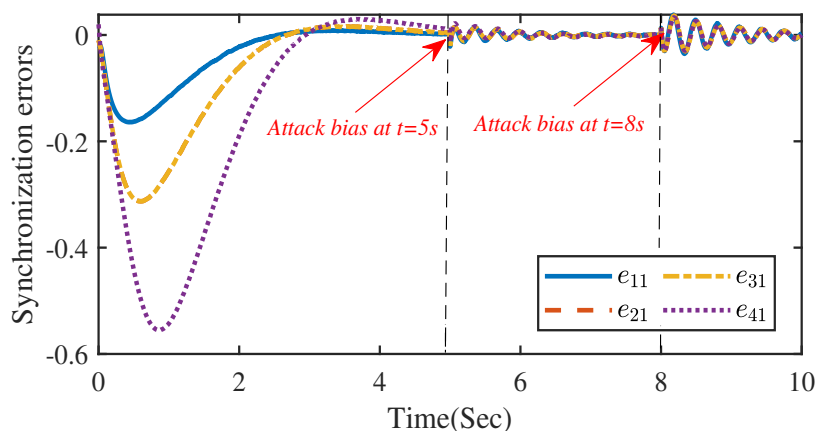


Figure 12. Synchronization errors of the four DGs.

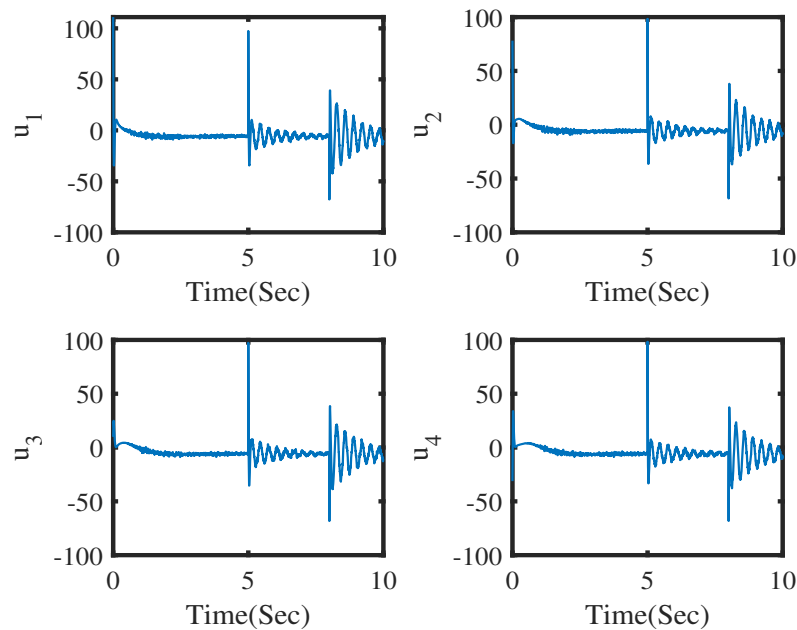


Figure 13. Voltage recovery control protocol.

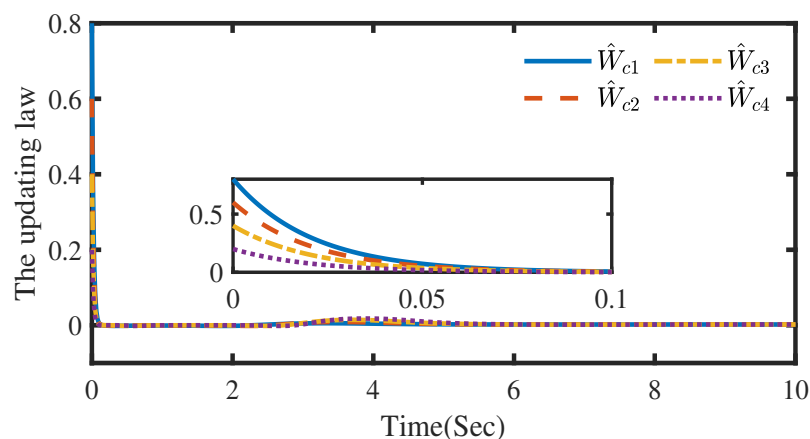


Figure 14. The updating laws of the four DGs.

The simulation results indicate that, even when subjected to sensor attacks, voltage output values of the DG can be observed, and they can accurately track the voltage reference value. Consequently, the designed optimal CSC approach effectively addresses the voltage recovery challenge in the island microgrids test system while ensuring the boundedness of all closed-loop signals.

6. Conclusions

This paper has addressed the optimal CSC problem for uncertain stochastic MASs over sensor attacks. A two-layer structure comprising an output feedback layer and a consensus error feedback layer has been constructed. This structural method has realized the transformation of the CSC problem

of stochastic MASs under output feedback form to the equivalent stochastic game problem of the consensus error dynamics. Subsequently, the ADP algorithm has been employed to formulate a stochastic game-based distributed consensus optimal control protocol, where a designed weight update law has been utilized for online estimation of the cost function. The constructed optimal CSC scheme guarantees both the minimization of the cost function and ensures the MS-UUB of all errors in the closed-loop system.

The proposed two-layer game-based framework can be extended along several directions: (i) the generalization to directed and switching, resource-constrained communication topologies, where the graphical-game equilibrium should be analyzed under time-varying interactions; (ii) the incorporation of more realistic attack classes, such as coordinated multi-point attacks and false-data-injection attacks with partial observability, which can be modeled within the same zero-sum-game framework by enlarging the adversary's strategy space; (iii) the development of a fully model-free policy-iteration algorithm so that the coupled HJI equations can be solved online without requiring an exact model of the system's nonlinearities (which is desirable in practical applications where $q_{i,s}$ and $g_{i,s}$ are unknown).

A number of practical issues should be highlighted when implementing the scheme: (i) the coupled HJI equations have to be solved iteratively, which demands an ADP-based value iteration that may be computationally demanding; (ii) the performance is sensitive to the selection of the critic-network basis functions and the design parameters, and these parameters should be tuned carefully (typically via offline pre-training or a standard trial-and-error procedure); (iii) the present analysis relies on the existence of a bounded admissible Nash solution and on the availability of the communication topology.

Author contributions

Zan Li: Conceptualization, methodology, writing—original draft. Kan He: Methodology. Zhongjun Yang: Analysis, validation. Lian Lian: Analysis, validation. Yifei Sun: Writing—review and editing. Huaici Zhao: Funding acquisition, writing—review and editing.

Use of Generative-AI tools declaration

The authors declare that they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This work is supported by the Central Government-Guided Local Science and Technology Development Fund (Grant no. 2026JH6/100400021), the 2026 Liaoning Provincial Major Science and Technology Special Project (Grant no. 2026JH1/11700024), the Liaoning Provincial Applied Basic Research Program (Grant no. 2026JH2/101300068), the Liaoning Provincial International Industrial Technology Research and Development Project (Grant no. 2026JH2/101900030), and the 2026 Basic Scientific Research Project for Universities in Liaoning Province (Grant no. LJ212610149013).

Conflict of interest

The author declares that there is no conflict of interest.

References

1. J. Wang, Z. Hu, H. Zhang, Multiple-AUV formation control with a collision avoidance algorithm based on a virtual structure and an improved artificial potential field method, *Meas. Sci. Technol.*, **36** (2025), 116203. <https://doi.org/10.1088/1361-6501/ae1b2b>
2. Y. Li, L. Meng, B. Dong, T. An, X. Cheng, S. Han, Friction disturbance model-based enhancement and optimization of active disturbance rejection controller for optoelectronic mechanical platforms, *Meas. Sci. Technol.*, **37** (2026), 156204. <https://doi.org/10.1088/1361-6501/ae5cc3>
3. H. Shan, Y. Jiang, H. Liang, S. Wang, H. K. Lam, Fuzzy adaptive containment control for nonlinear multi-manipulator systems with actuator faults and predefined accuracy, *IEEE Trans. Netw. Sci. Eng.*, **11** (2024), 1510–1523.
4. R. You, Q. Liu, S. Guo, Z. Cui, Neural observer-based iterative learning leader-following control for multi-agent systems subject to unknown dynamics and multiple faults, *Nonlinear Dyn.*, **114** (2026), 155.
5. G. Chen, X. Peng, Y. Lei, C. Huang, H. Li, Fast anomaly detection and joint state estimation for perturbed nonlinear systems with prolonged output anomalies, *IEEE/CAA J. Autom. Sinica*, **13** (2026), 1301–1313.
6. G. Chen, C. Huang, Y. Lei, X. Peng, H. Li, Fast attack detection and state estimation for nonlinear systems with vulnerable measurements: A fixed-time observer approach, *IEEE/CAA J. Autom. Sinica*, (2026). <https://doi.org/10.1109/JAS.2026.126311>
7. Y. Lei, Z. Wang, X. Wang, H. Li, I. C. Morarescu, Hybrid formation control for heterogeneous uncertain linear two-time-scale systems, *Automatica*, **176** (2025), 112267. <https://doi.org/10.1016/j.automatica.2025.112267>
8. B. Wen, J. Huang, X. Su, Y. Yang, Consensus control of nonlinear stochastic multiagent systems with unknown and time-varying control coefficients based on novel Nussbaum functions, *IEEE T. Syst. Man Cy-S.*, **55** (2025), 2264–2276. <https://doi.org/10.1109/TSMC.2024.3523369>
9. J. Chen, J. Shen, W. Chen, J. Li, S. Zhang, Application of robust fuzzy cooperative strategy in global consensus of stochastic multi-agent systems, *IEEE T. Autom. Sci. Eng.*, **22** (2025), 12058–12070. <https://doi.org/10.1109/TASE.2025.3539624>
10. Y. Li, J. H. Park, L. Wu, Q. Kong, Distributed output-feedback adaptive fuzzy leader-following consensus of stochastic nonlinear interconnected multiagent systems, *IEEE T. Syst. Man Cy-S.*, **52** (2022), 311–323.
11. H. Ren, H. Ma, H. Li, Z. Wang, Adaptive fixed-time control of nonlinear MASs with actuator faults, *IEEE/CAA J. Automatic.*, **10** (2023), 1252–1262.
12. X. Zhan, X. Du, J. Wu, H. Yan, Performance analysis of MIMO information time-delay systems with message queue under multiple communication parameters, *IEEE T. Syst. Man Cy-S.*, **54** (2024), 88–96. <https://doi.org/10.58250/jnanabha.2024.54111>

13. J. Ni, S. Zhao, J. Cao, Z. Wang, Predefined-time consensus tracking of high-order multiagent system with deception attack, *Inf. Sci.*, **649** (2023), 119671. <https://doi.org/10.1016/j.ins.2023.119671>
14. Z. Guo, X. Wang, M. Tian, J. Hu, Zonotopic state estimation and fault detection for discrete-time linear systems under DoS attacks: A switching controller design mechanism, *Automatica*, **185** (2026), 112833. <https://doi.org/10.1016/j.automatica.2026.112833>
15. X. Wang, Y. Cao, B. Niu, Y. Song, A novel bipartite consensus tracking control for multiagent systems under sensor deception attacks, *IEEE T. Cybernetics*, **53** (2023), 5984–5993. <https://doi.org/10.1109/TCYB.2022.3225361>
16. Z. Guo, Q. Wei, X. Zhao, B. Wang, B. Niu, H. Liu, Attack detection and active attack defense for cyber–physical systems via zonotopic observer and reachability analysis, *IEEE T. Cybernetics*, <https://doi.org/10.1109/TCYB.2026.3670840>
17. Y. Yang, Y. Xiao, T. Li, Attacks on formation control for multiagent systems, *IEEE T. Cybernetics*, **52** (2022), 12805–12817. <https://doi.org/10.1109/TCYB.2021.3089375>
18. Z. Wu, N. Zhao, A. A. Albeshri, Deferred actuator replacement-based performance recovery consensus control for nonlinear multi-agent systems under DoS attacks, *Nonlinear Dynam.*, **114** (2026), 865.
19. N. Xu, F. Wei, G. Zong, H. Wang, X. Zhao, Adaptive prescribed performance secure consensus control for high-order constrained multi-agent systems with time-varying powers and arbitrary initial states, *IEEE T. Autom. Sci. Eng.*, **23** (2026), 9134–9145. <https://doi.org/10.1109/TASE.2026.3686874>
20. L. Chen, S. Tong, Observer-based adaptive fuzzy consensus control of nonlinear multiagent systems encountering deception attacks, *IEEE T. Ind. Inform.*, **20** (2024), 1808–1818. <https://doi.org/10.1109/TII.2023.3281703>
21. H. Ren, Z. Cheng, J. Qin, R. Lu, Deception attacks on event-triggered distributed consensus estimation for nonlinear systems, *Automatica*, **154** (2023), 111100. <https://doi.org/10.1016/j.automatica.2023.111100>
22. H. Zhang, X. Guo, J. Sun, Y. Zhou, Event-triggered cooperative adaptive fuzzy control for stochastic nonlinear systems with measurement sensitivity and deception attacks, *IEEE T. Fuzzy Syst.*, **31** (2022), 774–785. <https://doi.org/10.1071/WF22006>
23. W. He, W. Xu, X. Ge, Q. L. Han, W. Du, F. Qian, Secure control of multiagent systems against malicious attacks: A brief survey, *IEEE T. Ind. Inform.*, **18** (2021), 3595–3608. <https://doi.org/10.1109/TII.2021.3126644>
24. X. Gao, F. Deng, P. Zeng, Zero-sum game-based security control of unknown nonlinear Markov jump systems under false data injection attacks, *Int. J. Robust Nonlinear Control*, (2022), In press.
25. C. Wu, X. Li, W. Pan, J. Liu, L. Wu, Zero-sum game-based optimal secure control under actuator attacks, *IEEE T. Autom. Control*, **66** (2021), 3773–3780. <https://doi.org/10.1109/TAC.2020.3029342>
26. S. Gao, H. Zhang, Z. Wang, C. Huang, H. Yan, Optimal injection attack strategy for cyber-physical systems under resource constraint: A game approach, *IEEE T. Control Netw.*, **10** (2023), 636–646.

27. Q. Wei, R. Song, P. Yan, Data-driven zero-sum neuro-optimal control for a class of continuous-time unknown nonlinear systems with disturbance using ADP, *IEEE T. Neural Net. Lear.*, **27** (2016), 444–458. https://doi.org/10.1111/cnr.456_12958
28. T. An, B. Dong, H. Yan, L. Liu, B. Ma, Dynamic event-triggered strategy-based optimal control of modular robot manipulator: A multiplayer nonzero-sum game perspective, *IEEE T. Cybernetics*, **54** (2024), 7514–7526. <https://doi.org/10.1109/TCYB.2024.3468875>
29. T. An, Y. Wang, G. Liu, Y. Li, B. Dong, Cooperative game-based approximate optimal control of modular robot manipulators for human-robot collaboration, *IEEE T. Cybernetics*, **53** (2023), 4691–4703. <https://doi.org/10.1109/TCYB.2023.3277558>
30. H. Jiang, H. Zhang, X. Xie, Critic-only adaptive dynamic programming algorithms applications to the secure control of cyber–physical systems, *ISA T.*, **104** (2020), 138–144. <https://doi.org/10.1016/j.isatra.2019.02.012>
31. L. Zhang, Y. Chen, M. Li, ADP-based remote secure control for networked control systems under unknown nonlinear attacks in sensors and actuators, *IEEE T. Ind. Inform.*, **18** (2022), 6003–6014. <https://doi.org/10.1109/TII.2021.3137816>
32. F. Wang, B. Chen, C. Lin, X. Li, Distributed adaptive neural control for stochastic nonlinear multiagent systems, *IEEE T. Cybernetics*, **47** (2017), 1795–1803. <https://doi.org/10.1109/TCYB.2016.2623898>
33. Y. Yan, T. Li, J. Wang, C. P. Chen, H. Liang, Fuzzy prescribed-time self-triggered consensus control for nonlinear multi-agent systems with dead-zone output, *Fuzzy Sets Syst.*, **492** (2024), 109071. <https://doi.org/10.1016/j.fss.2024.109071>
34. D. Wang, C. Mu, Adaptive-critic-based robust trajectory tracking of uncertain dynamics and its application to a spring–mass–damper system, *IEEE T. Ind. Electron.*, **65** (2018), 654–663. <https://doi.org/10.1109/TIE.2017.2722424>
35. J. Jeong, Y. Lim, A. Parivallal, An asymmetric Lyapunov-Krasovskii functional approach for event-triggered consensus of multi-agent systems with deception attacks, *Appl. Math. Comput.*, **439** (2023), 127584. <https://doi.org/10.1016/j.amc.2022.127584>
36. G. Liu, Q. Sun, R. Wang, X. Hu, Nonzero-sum game-based voltage recovery consensus optimal control for nonlinear microgrids system, *IEEE T. Neural Netw. Lear.*, **34** (2023), 8617–8629. <https://doi.org/10.1109/TNNLS.2022.3151650>



AIMS Press

© 2026 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>)