An Improved Bayesian Graphical Game Method for the Optimal Consensus Problem in the Presence of False Information

Yue Zhang, Yan-Wu Wang, *Senior Member, IEEE*, Yan Lei, Yun-Feng Luo, and Zhi-Wei Liu, *Senior Member, IEEE*

***Abstract*—When realizing multiagent optimal consensus control, it may encounter the situation that malicious agents transmit false information. Besides, due to the unreliability of information interaction and the uncertainty of the system itself, the agent may not fully know its own cost function. In this article, an improved Bayesian graphical game method is proposed to solve the optimal consensus problem of linear dynamical networks in the presence of false information. The agent's probabilistic estimate of the uncertainty is named belief, and the update of probability estimate is named belief update. An information tradeoff principle is designed to solve the belief update problem in the presence of malicious neighbors. The principle can not only reduce the loss caused by false information but also force malicious agents to switch from deception to cooperation. On this basis, a new belief update method with information tradeoff principle is established. It is proved that this new belief update method has a faster convergence rate than the Bayesian belief update method when malicious neighbors exist. As an illustrative example, the graphical game solution is applied to the formation tracking control problem of a quad-rotor unmanned aerial vehicle swarm. Theoretical proof and simulation comparisons can illustrate the proposed method's advantages.**

***Index Terms*—Bayesian graphical game, belief update method, information tradeoff principle, optimal consensus control.**

I. Introduction

In recent years, multiagent systems (MASs) have been widely applied in smart grids [1], autonomous underwater vehicles [2], persistent monitoring [3], and so on. In the study of MAS, the consensus control problem has always been a hot topic. The consensus control problem requires each agent to be equipped with a certain control protocol based on local information so that the states of all agents reach a common value. Consensus control has attracted considerable attention in many areas, such as manipulator systems [4], networked quadrotor aircraft [5], and so on.

Since a single agent cannot directly obtain all the information of the whole system, the realization of consensus control depends on the information exchange between agents. However, false information may exist in the communication. False information generally comes from false data injection [6] or deception attacks by malicious neighbors [7]. Although there have been many studies on consensus control in the presence of false information, most of the literature does not consider that the agent may have its own cost function that needs to be optimized.

The cost functions of the agents are generally related to both itself and its neighbors [8], [9], [10], [11], [12], [13]. It requires the agent to consider not only its own behavior but also the behavior of neighbor agents. However, due to the unreliability of information interaction and the uncertainty of the system matrices, the agents may not fully know about their own cost functions. It leads to an optimal consensus problem with incomplete information. Lopez et al. [11] used differential graphical game theory to solve the optimal consensus problem in cases where agents do not fully know their own cost functions. They build a Bayesian graphical game, revealing the relationship between beliefs and distributed optimal strategies under incomplete information conditions. Additionally, this article discusses the convergence of two belief update methods: 1) Bayesian and 2) non-Bayesian. Lopez et al. [11] also demonstrated that solving Bayesian graphical games relies on solving coupled Bayesian Hamilton Jacobi (BHJ) equations. As the simulation in [11] showed, the process of solving the BHJ equations in [11] requires the state information of not only the agents' neighbors, but also their neighbors' neighbors. This results in nondistributed solutions.

This article uses the Bayesian graphical game method to solve the optimal consensus problem with incomplete information. Inspired by [12], a modified cost function different from [11] is adopted to decouple the BHJ equations and ensure the existence of distributed solutions, which only depend on local information of their own and their immediate neighbors on the communication graph. On this basis, an improved Bayesian graphical game method is proposed to solve the optimal consensus problem in the presence of false information.

Manuscript received 17 September 2023; accepted 17 February 2024. Date of publication 6 March 2024; date of current version 17 May 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 62233006, Grant 62173152, Grant 62103156, and Grant 62303189. This article was recommended by Associate Editor L. C. Rego. *(Corresponding author: Yan-Wu Wang.)*

The authors are with the Key Laboratory of Image Processing and Intelligent Control and the School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: wangyw@hust.edu.cn).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TSMC.2024.3367961.

Digital Object Identifier 10.1109/TSMC.2024.3367961

In MAS Bayesian game, belief update is a critical step [14], [15], [16]. In an uncertain environment, collecting environmental information and modifying the original cognition after collecting new information is very important for the agent. This process is usually defined as the belief update, changing the original belief and accepting more accurate and reliable information that may be inconsistent with the original belief. The Bayesian update method in [11] is similar to the distributed virtual game algorithm in [17], where agents observe the behavior of neighbors, share information with other agents, learn unknown information through interactive processes, and use expected utility functions before fully understanding environmental information. In the process of belief update, malicious agents may transmit false information [18]. In scenarios of incomplete information games, the presence of false information can potentially lead to catastrophic consequences. For instance, false data injection attacks in power systems can cause bus voltage instability and cascading failures, while false beliefs about neighboring parameters in intelligent vehicles can result in vehicle collisions in intersections [19], [20]. Therefore, it is highly valuable to consider how to address the potential existence of false information during the process of belief update when studying Bayesian graphical game problems. However, Lopez et al. [11] did not consider the problem that neighbor agents may transmit false information when solving the Bayesian graphical game.

This article considers the situation where there exist malicious agents during the process of belief update. Malicious agents may transmit false information during the process of belief update because they fear that their privacy may be exposed or their costs may increase. But a malicious agent wants to use the true information of its neighbors. If agents continue to use the Bayesian belief update method in [11], they cannot distinguish false information and they need to discard the information from all their neighbors and only use normal estimation to predict unknown information. As presented in [11], when using normal estimation, the agent models the probability distribution of unknown information with a normal distribution and adopts the normal mean as the expected value of the unknown information. The normal estimation strategy is somehow conservative and results in slow convergence of the consensus. In this article, we design an information tradeoff principle to fully use the information in the graph. We then design a new belief update method with information tradeoff principle and prove that this new belief update method has a faster convergence rate than the Bayesian belief update method when malicious neighbors exist.

Potential applications for the proposed Bayesian graphical game include rational and persistent deception among intelligent robots, resilient control of multimicrogrids against false data injection attacks, and sensible decision-making against deception attacks by possibly malicious power nodes [21]. As the number of microgrids continues to increase in urban areas, the development of optimal strategies for unknown scenarios becomes increasingly necessary.

The contributions of this article are the following.

1) Inspired by [12], a new cost function is chosen to avoid the problem in [11] that the BHJ equations do not have distributed solutions. On this basis, an improved Bayesian graphical game method is used to solve the optimal consensus problem in the presence of false information. The cost function in [12] can only provide distributed solutions for the complete information game while the cost function in this article is used to provide distributed Nash solutions for the incomplete information game. So the two cost functions are not identical.
2) An effective information tradeoff principle is established to deal with the problem that there exist malicious neighbors during the process of belief update. The information tradeoff principle can not only reduce the loss caused by false information but also force malicious agents to switch from deception to cooperation, that is, to transmit true information.
3) A new belief update method with information tradeoff principle is established, and it is proved that the new belief update method has a faster convergence rate than the Bayesian belief update method when malicious neighbors exist.

The structure of this article is as follows. In Section II, the optimal consensus problem with incomplete information is proposed, and the optimal strategy is obtained. In Section III, it is proved theoretically that the new belief update method with the information tradeoff principle has a faster convergence rate. Section IV proves that the global error dynamics are asymptotically stable, and the optimal strategies of all agents will form a Nash equilibrium. Section V presents simulation examples and comparison studies to show the method's effectiveness. Concluding remarks and possible future work are discussed in Section VI.

Notation: Define $U = \sqrt{V}$ if $V = U^T U$. I is an identity matrix. $A > (\geq 0)$ stands for positive definite (semidefinite) matrix. $A \otimes B$ is the Kronecker product between the matrices A and B.

II. Bayesian Graphical Game for the Optimal Consensus Problem

Consider a MAS with N agents connected by a communication graph $\mathcal{G} = (V, \mathcal{E}, E)$. $V = \{1, \ldots, N\}$ serves as the agents set. $\mathcal{E} \subseteq V \times V$ is the set of edges. $E = [a_{ij}] \in R^{N \times N}$ is the adjacent matrix, where a_{ij} is the edge weights of the communication graph. If $(j, i) \in \mathcal{E}$, $a_{ij} > 0$; otherwise, $a_{ij} = 0$. The set of neighbors of node i is $\mathcal{N}_i = \{j : a_{ij} > 0\}$. Let the number of elements in $\mathcal{N}_i$ be N_i. g_i is the pinning gain of node i. If node i is contained by the leader, $g_i > 0$; otherwise, $g_i = 0$. The weighted in-degree of i is $d_i = \sum_{j=1}^{N} a_{ij}$. The in-degree matrix is defined as $D = \mathrm{diag}(d_1, d_2, \ldots, d_N)$ and Laplacian matrix is defined as $L = D - E$. Denote the pining matrix as $G = \mathrm{diag}(g_1, g_2, \ldots, g_N)$.

The linear dynamics of the N agents in the graph are described by

$$\dot{x}_i = Ax_i + Bu_i, \; i = 1, \ldots, N \tag{1}$$

where $x_i(t) \in R^n$ is the state variable of agent i and $u_i \in R^{m_i}$ is the control input of agent i. A is the state matrix. B is the control input matrix.

Suppose the virtual leader is described by

$$\dot{x_0} = Ax_0 \tag{2}$$

where $x_0(t) \in R^n$ is the state variable of the leader. The MAS consensus requires that

$$\lim_{t\to\infty} \|x_i(t) - x_j(t)\| = 0 \quad \forall i, j = 1, \dots, N \tag{3}$$

$$\lim_{t\to\infty} \|x_i(t) - x_0(t)\| = 0 \quad \forall i = 1, \dots, N \tag{4}$$

hold for any initial state $x_i(0), i = 1, \dots, N$.

As pointed out in [22], these homogeneous dynamics (i.e., A and B matrices in (1) are the same) are beneficial for achieving the matching of aerial vehicles' velocities in simulation.

Assumption 1: The pair (A, B) is controllable.

Assumption 2: The graph $\mathcal{G}$ is strongly connected and at least one agent can receive information from the leader.

The local error of agent i is defined as

$$\delta_i = \sum_{j=1}^{N} a_{ij}(x_i - x_j) + g_i(x_i - x_0). \tag{5}$$

In achieving consensus, each agent hopes to minimize its cost. So a multiplayer game problem is formed.

Similar to [11], define the type space as a finite set to describe different intentions of the agents. For example, in the consensus problem for frequency and voltage restoration of microgrids, the different degrees of actuation/propulsion faults and parameter disturbances will affect the control effort required for each node to achieve consensus. We define the degree of faults and parameter perturbations for each node as the node's type. Thus, the degrees of faults and parameter perturbations for all nodes constitute the global type of microgrid systems. Each node corresponds to different control cost functions under different global types. Since the degrees of faults and parameter perturbations are typically uncertain, the type of each node is random. Node i infers the probability distribution of its neighbor's types based on historical information, such as voltage, frequency, fault degree, and information, obtained through communication interactions with its neighbors, subsequently deducing the true global type. Agent i knows all the possibilities of its neighbors' type. The type combination of the entire system is defined as the global type. Let $\Theta = \Theta_1 \times \cdots \times \Theta_N$, where Θ is the global type space and Θ_i is the type space of agent i. $\theta = \theta_1 \times \cdots \times \theta_N$ is the global type, $\theta \in \Theta$, $\theta_i \in \Theta_i$. θ_i is one of n_i types $\theta_i^1, \theta_i^2, \dots, \theta_i^{n_i}$. Different type $\theta \in \Theta$ corresponds to different cost function, i.e.,

$$J_i^\theta = \int_0^\infty \sum_{j\in\mathcal{N}_i} \Big(\delta_{ij}^T Q_{ij}^\theta \delta_{ij} + u_i^T R_{ii}^\theta u_i - a_{ij} u_j^T R_{ij}^\theta u_j\Big) dt \tag{6}$$

where $\delta_{ij} = [\delta_i^T, \delta_j^T]^T$, and $Q_{ij}^\theta = [Q_i^\theta, \tilde{Q}_{ij}^\theta; \tilde{Q}_{ij}^{\theta\mathrm{T}}, \hat{Q}_{ij}^\theta]$. $R_{ii}^\theta > 0$, $Q_i^\theta \geq 0$, $Q_{ij}^\theta \geq 0$, and $R_{ij}^\theta \geq 0$. $Q_i^\theta \geq 0$ means that agent i tends to minimize its own local error δ_i. The terms $2\delta_i^T \tilde{Q}_{ij}^\theta \delta_j$ and $\delta_j^T \hat{Q}_{ij}^\theta \delta_j$ mean agent i also cares about the local error of its neighbors. $u_i^T R_{ii}^\theta u_i$ is used to minimize the control effort of i while $-a_{ij} u_j^T R_{ij}^\theta u_j$ is used to maximize the control effort of j. In general, when accomplishing a common task goal together, agent i hopes that the entire group can achieve the goal at the minimum cost. However, during the allocation of this minimum cost, agent i would like to allocate less compared to its neighbors, essentially trying to complete the task more "lazily."

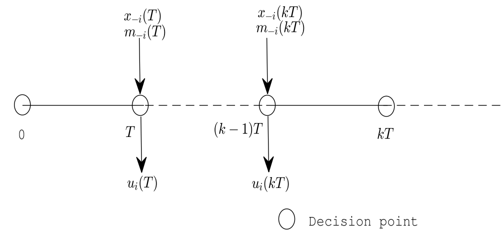

Fig. 1. Communication mechanism.

Remark 1: Note that the types of agents are only related to the cost functions but not to whether an agent is malicious or not. In this article, whether an agent is malicious is only to classify its behavior when transferring information to its neighbors. The cheating behavior of a malicious agent will lead to the presence of false information. The types of agents are unknown but constant information, while the maliciousness of an agent is unknown and dynamic information.

Remark 2: The inspiration for the design of (6) comes from the cost function under the condition of complete information in [12]. Comparing (6) with the existing cost function under the condition of incomplete information in [11], two extra terms, i.e., $2\delta_i^T \tilde{Q}_{ij}^\theta \delta_j$ and $\delta_j^T \hat{Q}_{ij}^\theta \delta_j$ are introduced. The two extra terms are helpful to decouple the BHJ equations and to ensure that there exist distributed solutions to the BHJ equations. We will discuss their role in detail in Remark 3.

In order to make the best decisions in the game, each agent wants to know its actual cost function. The actual cost function is determined by the actual global type θ^*. Because each agent i only knows its own type θ_i but does not know the type of other agents, it will use the obtained information to estimate the global type. There are two different kinds of information that an agent can obtain. One is the neighbors' historical states information of the previous sampling instants, the other is the states information that the neighbors currently claim to be at the next sampling instant. The neighbors' historical states observed by the agent are true and reliable while the states information claimed by the neighbors may be false.

The value function of each agent i is defined as

$$V_i^\theta(\delta_i, \delta_{-i}) = \int_t^\infty \sum_{j\in\mathcal{N}_i} \Big(\delta_{ij}^T Q_{ij}^\theta \delta_{ij} + u_i^T R_{ii}^\theta u_i - a_{ij} u_j^T R_{ij}^\theta u_j\Big) d\tau.$$

Let x_{-i} be the states set of i's neighbors and m_{-i} be the states set claimed by i's neighbors. As shown in Fig. 1, each agent designs its own control strategy after obtaining information x_{-i} and m_{-i} from neighbors. Let $p(\theta \mid \theta_i, m_{-i})$ be the probability that agent i estimates the global type to be θ based on θ_i and m_{-i}.

Before the estimation probability of the actual global type θ^* reaches 1, agent i uses the expected cost function to design its own control strategy. The expected cost function is

$$\mathbb{E}J_i(\delta_i, \delta_{-i}, u_i, u_{-i}) = \sum_{\theta\in\Theta} p(\theta \mid \theta_i, m_{-i}) J_i^\theta(\delta_i, \delta_{-i}, u_i, u_{-i})$$
$$= \mathbb{E}V_i(\delta_i(0), \delta_{-i}(0)) = \sum_{\theta\in\Theta} p(\theta \mid \theta_i, m_{-i}) V_i^\theta(\delta_i(0), \delta_{-i}(0)).$$

In achieving consensus, agents wish that the optimal control strategy $u_i^*, i \in V$, satisfies

$$\mathbb{E}J_i^* \triangleq \mathbb{E}J_i\left(\delta_i, \delta_{-i}, u_i^*, u_{\mathcal{G}-i}^*\right) \le \mathbb{E}J_i\left(\delta_i, \delta_{-i}, u_i, u_{\mathcal{G}-i}^*\right)$$

where $\mathcal{G} - i$ represents all the agents in the graph except i.

Let

$$\tilde{V}_i^\theta\left(\delta_i, \delta_j\right) = \int_t^\infty \left(\delta_{ij}^T Q_{ij}^\theta \delta_{ij} + u_i^T R_{ii}^\theta u_i - a_{ij} u_j^T R_{ij}^\theta u_j\right) d\tau.$$

Obviously $V_i^\theta(\delta_i, \delta_{-i}) = \sum_{j\in\mathcal{N}_i} \tilde{V}_i^\theta(\delta_i, \delta_j)$.

Referencing [23], define the expected Hamiltonian as

$$\mathbb{E}H_i(\delta_i, u_i, u_{-i}, \theta) = \sum_{\theta\in\Theta} p(\theta \mid \theta_i, m_{-i}) \times \sum_{j\in\mathcal{N}_i}\left(\frac{\partial \tilde{V}_i^{\theta\mathrm{T}}}{\partial \delta_{ij}}\dot{\delta}_{ij} + \delta_{ij}^T Q_{ij}^\theta \delta_{ij} + u_i^T R_{ii}^\theta u_i - a_{ij} u_j^T R_{ij}^\theta u_j\right). \quad (7)$$

Referencing [12], to guarantee a distributed solution, $\tilde{V}_i^\theta(\delta_i, \delta_j)$ is required to depend on δ_i solely, i.e., $\tilde{V}_i^\theta(\delta_i, \delta_j) = \tilde{V}_i^\theta(\delta_i)$. Then

$$\frac{\partial \tilde{V}_i^{\theta\mathrm{T}}}{\partial \delta_{ij}}\dot{\delta}_{ij} = \begin{bmatrix} \frac{\partial \tilde{V}_i^{\theta\mathrm{T}}}{\partial \delta_i} & \frac{\partial \tilde{V}_i^{\theta\mathrm{T}}}{\partial \delta_j}\end{bmatrix}\begin{bmatrix}\dot{\delta}_i \\ \dot{\delta}_j\end{bmatrix} = \frac{\partial \tilde{V}_i^{\theta\mathrm{T}}}{\partial \delta_i}\dot{\delta}_i$$
$$= \frac{\partial \tilde{V}_i^{\theta\mathrm{T}}}{\partial \delta_i}\left(A\delta_i + (d_i + g_i)Bu_i - \sum_{j\in\mathcal{N}_i} a_{ij}Bu_j\right).$$

Then (7) is employed to obtain the optimal strategy by computing its derivative with respect to u_i and equating it to zero, i.e.,

$$\sum_{\theta\in\Theta} p(\theta \mid \theta_i, m_{-i}) \sum_{j\in\mathcal{N}_i}\left((d_i + g_i)B^T\frac{\partial \tilde{V}_i^\theta}{\partial \delta_i} + 2R_{ii}^\theta u_i\right) = 0. \quad (8)$$

Solving (8) yields the optimal strategy

$$u_i^* = -\frac{1}{2}(d_i + g_i) \times \left[\sum_{\theta\in\Theta} p(\theta \mid \theta_i, m_{-i}) R_{ii}^\theta\right]^{-1} \sum_{\theta\in\Theta} p(\theta \mid \theta_i, m_{-i}) B^{\mathrm{T}} \frac{\partial \tilde{V}_i^\theta}{\partial \delta_i}. \quad (9)$$

Let $\mathbb{E}Q_i = \sum_{\theta\in\Theta} p\ (\theta|\theta_i, m_{-i})Q_i^\theta$, $\mathbb{E}\tilde{Q}_{ij} = \sum_{\theta\in\Theta} p\ (\theta|\theta_i, m_{-i})\ \tilde{Q}_{ij}^\theta$, $\mathbb{E}\hat{Q}_{ij} = \sum_{\theta\in\Theta} p\ (\theta|\theta_i, m_{-i})\ \hat{Q}_{ij}^\theta$, $\mathbb{E}R_{ii} = \sum_{\theta\in\Theta} p\ (\theta|\theta_i, m_{-i})R_{ii}^\theta$, $\mathbb{E}R_{ij} = \sum_{\theta\in\Theta} p\ (\theta|\theta_i, m_{-i})R_{ij}^\theta$, and $\mathbb{E}\,\tilde{V}_i^{\cdot} = \sum_{\theta\in\Theta} p\ (\theta|\theta_i, m_{-i})\ ([\partial \tilde{V}_i^\theta]/\ [\partial\delta_i])$. Substituting (9) into (7), then $\tilde{V}_i^\theta$ is the solution of the following BHJ equation:

$$\sum_{\theta\in\Theta} p(\theta \mid \theta_i, m_{-i}) \sum_{j\in\mathcal{N}_i}\Bigg(\delta_i^T Q_i^\theta \delta_i + \frac{\partial \tilde{V}_i^{\theta\mathrm{T}}}{\partial \delta_i}A\delta_i - \frac{(d_i+g_i)^2}{4}\frac{\partial \tilde{V}_i^{\theta\mathrm{T}}}{\partial \delta_i}B(\mathbb{E}R_{ii})^{-1}\cdot B^T\left(\mathbb{E}\dot{\tilde{V}}_i\right) + 2\delta_i^T\tilde{Q}_{ij}^\theta\delta_j$$
$$+ \frac{1}{2}\sum_{j\in\mathcal{N}_j} a_{ij}\left(d_j + g_j\right)\frac{\partial \tilde{V}_i^{\theta\mathrm{T}}}{\partial \delta_i}B\left(\mathbb{E}R_{jj}\right)^{-1}\cdot B^T\left(\mathbb{E}\dot{\tilde{V}}_j\right) + \delta_j^T\hat{Q}_{ij}^\theta\delta_j$$
$$- \frac{a_{ij}}{4}\left(d_j + g_j\right)^2\left(\mathbb{E}\dot{\tilde{V}}_j\right)^T B\left(\mathbb{E}R_{jj}\right)^{-1}R_{ij}^\theta\left(\mathbb{E}R_{jj}\right)^{-1}B^T\left(\mathbb{E}\dot{\tilde{V}}_j\right)\Bigg) = 0. \quad (10)$$

Inspired by [12], we give Lemma 1 and Theorem 1 to ensure the existence of the distributed solution $\tilde{V}_i^\theta(\delta_i)$ to (10).

Lemma 1: If there exists a distributed solution in the form of

$$\tilde{V}_i^\theta(\delta_i) = \delta_i^{\mathrm{T}}\tilde{P}_i^\theta\delta_i \quad (11)$$

to the BHJ equation (10), where $\tilde{P}_i^\theta > 0$, then the optimal strategy of agent i is unique and takes the form

$$u_i^* = -(d_i + g_i) \times \left[\sum_{\theta\in\Theta} p(\theta \mid \theta_i, m_{-i}) R_{ii}^\theta\right]^{-1} \sum_{\theta\in\Theta} p(\theta \mid \theta_i, m_{-i}) B^T\tilde{P}_i^\theta\delta_i. \quad (12)$$

Proof: Please see the Appendix-A.

Remark 3: Based on (22)–(24) in [11] and its statements in the first paragraph of Section III-D, it is evident that the solution to the BHJ equation in [11] relies on the state information of neighbors' neighbors. When solving the BHJ equation for agent i in [11] to obtain its optimal controller u_i and control matrix P_i, nonlocal information δ_j is required. However, in this article, two extra terms in (6) are helpful to decouple (10) so that we can obtain the distributed solution $\tilde{V}_i^\theta(\delta_i)$ to (10) and optimal controller u_i by (13), which does not depend on δ_j. From Theorem 1, it can be seen that selecting the appropriate extra terms can cut the coupling terms in (10) as seen from (24). As a result, the improved Bayesian graphical game method in this article can permit a distributed solution.

Theorem 1: Let $\tilde{V}_i^\theta(\delta_i) = \delta_i^T\tilde{P}_i^\theta\delta_i$, $\mathbb{E}\tilde{P}_i = \sum_{\theta\in\Theta} p(\theta \mid \theta_i, m_{-i})\tilde{P}_i^\theta$. If $(A, \sqrt{\sum_{\theta\in\Theta} p(\theta \mid \theta_i, m_{-i})Q_i^\theta})$ is observable and (A, B) is stabilizable, there exists a unique positive definite solution $\mathbb{E}\tilde{P}_i$ to

$$\left(\mathbb{E}\tilde{P}_i\right)^T A + A^T\left(\mathbb{E}\tilde{P}_i\right) + \mathbb{E}Q_i - (d_i + g_i)^2\left(\mathbb{E}\tilde{P}_i\right)^T B(\mathbb{E}R_{ii})^{-1}B^T\mathbb{E}\tilde{P}_i = 0. \quad (13)$$

Then, the coupled BHJ equation (10) holds for all possible δ_i and δ_j, if $\mathbb{E}\tilde{Q}_{ij}$ and $\mathbb{E}\hat{Q}_{ij}$ are chosen as

$$\mathbb{E}\tilde{Q}_{ij} = -a_{ij}\left(d_j + g_j\right)\left(\mathbb{E}\tilde{P}_i\right)^T\cdot N_iB\left(\mathbb{E}R_{jj}\right)^{-1}B^T\mathbb{E}\tilde{P}_j \quad (14)$$

$$\mathbb{E}\hat{Q}_{ij} = a_{ij}\left(d_j + g_j\right)^2\left(\mathbb{E}\tilde{P}_j\right)^T\cdot B\left(\mathbb{E}R_{jj}\right)^{-1}\left(\mathbb{E}R_{ij}\right)\left(\mathbb{E}R_{jj}\right)^{-1}B^T\mathbb{E}\tilde{P}_j. \quad (15)$$

Proof: Please see the Appendix-B.

In this article, for the optimal strategy (9), there is cognitive information $p(\theta|\theta_i, m_{-i})$ that needs to be updated, which is called the belief of agent i. Hence, (9) is only the optimal control strategy corresponding to the expected cost, and the agent is eager to know the actual global type θ^* so that they can use the actual cost function to calculate the true optimal control strategy. So the agents' goal is to get $p(\theta^*|\theta_i, m_{-i})$ close to 1 through certain belief update method.

III. Belief Update Method With Information Tradeoff Principle

In this section, we propose a belief update method with a new information tradeoff principle to deal with the problem of false information during the process of belief update. It can speed up the belief update compared with the Bayesian belief update.

A. Type Estimation With Information Transmission

In the process of belief update, let each agent update its belief at every sampling instant. The traditional Bayesian belief update used in [11] is

$$p(\theta|x_{-i}(t+T), x_{-i}(t), \theta_i) = \frac{p(x_{-i}(t+T)|x_{-i}(t), \theta)p(\theta|x_{-i}(t), \theta_i)}{p(x_{-i}(t+T)|x_{-i}(t), \theta_i)}. \tag{16}$$

In order to speed up the convergence of agent i' belief to the actual type θ^*, we introduce the concept of the signal game [24] in the process of belief update. The signal game is a game in which there are two players, the signal sender sending private information and the signal receiver making decisions on the basis of information transmitted from the signal sender.

Based on the signal game, we establish a new belief update method described as follows:

$$p(\theta|x_{-i}(t+T), x_{-i}(t), \theta_i, m_{-i}(t, t+T)) = \frac{p(x_{-i}(t+T)|x_{-i}(t), \theta, m_{-i}(t, t+T))p(\theta|x_{-i}(t), \theta_i)}{p(x_{-i}(t+T)|x_{-i}(t), \theta_i)} \tag{17}$$

where $m_{-i}(t, t+T)$ is the state information of time $t+T$ transmitted from i's neighbors $-i$ to agent i at the sampling time t. The information informs agent i what states $-i$ will be at $t+T$. But the information may be false. $p(\theta|x_{-i}(t+T), x_{-i}(t), \theta_i, m_{-i}(t, t+T))$ is agent i's belief about type θ at $t+T$, and $p(\theta|x_{-i}(t), \theta_i)$ is agent i's belief about type θ at t. $p(x_{-i}(t+T)|x_{-i}(t), \theta, m_{-i}(t, t+T))$ is the probability that i's neighbors $-i$ will reach $x_{-i}(t+T)$ at $t+T$ under the condition that its neighbors $-i$ reach the state $x_{-i}(t)$ at t in the given global type θ and the neighbors' information is $m_{-i}(t, t+T)$. $p(x_{-i}(t+T)|x_{-i}(t), \theta_i)$ is the probability that agent i's neighbors $-i$ reach $x_{-i}(t+T)$ at $t+T$ under the condition that agent i does not know the global type but know its neighbors $-i$ reach $x_{-i}(t)$ at t.

We can extend (17) to the kth round to get the following equation:

$$\begin{aligned} &p\begin{pmatrix} \theta|x_{-i}(t+kT), x_{-i}(t+(k-1)T), \\ \theta_i, m_{-i}(t+(k-1)T, t+kT) \end{pmatrix} \\ &= \frac{\prod_{j\in N_i} p\begin{pmatrix} x_j(t+kT)|x_{-i}(t+(k-1)T), \\ \theta, m_j(t+(k-1)T, t+kT) \end{pmatrix}}{p(x_{-i}(t+kT)|x_{-i}(t+(k-1)T), \theta_i)} \\ &\quad \times p(\theta|x_{-i}(t+(k-1)T), \theta_i), k = 1, 2, 3, \ldots \end{aligned} \tag{18}$$

Similarly we can extend the traditional Bayesian belief update (16) to the kth round to get the following equation:

$$\begin{aligned} &p(\theta|x_{-i}(t+kT), x_{-i}(t+(k-1)T), \theta_i) \\ &= \frac{\prod_{j\in N_i} p(x_j(t+kT)|x_{-i}(t+(k-1)T), \theta)}{p(x_{-i}(t+kT)|x_{-i}(t+(k-1)T), \theta_i)} \\ &\quad \times p(\theta|x_{-i}(t+(k-1)T), \theta_i), k = 1, 2, 3, \ldots. \end{aligned} \tag{19}$$

Remark 4: Comparing with the traditional Bayesian belief update method in [11], the difference of the kth round of belief update is that the information of $m_{-i}(t+(k-1)T, t+kT)$ is used when agents predict the value of $x_j(t+kT)$. This enables each agent to have a more informative way to understand the global type, so as to accelerate the convergence rate of the belief update. In order to ensure that each agent will not make an unreasonable strategy due to false or wrong information, agents must have a reasonable tradeoff principle to make a sound judgment on the information from the neighbors. For this reason, we propose an information tradeoff principle in the next section.

Remark 5: In the presence of false information, using the traditional Bayesian belief update means that agents will discard all the information from their neighbors without any judgment and only use the neighbors' historical states for estimation. Then the agent cannot fully use the information and can only update its belief based on local information. It results in a waste of information and makes the agent update at a lower rate. The information tradeoff principle in the next section can appropriately use this information.

B. Information Tradeoff Principle

In the MAS, there are two kinds of agents: 1) friendly agents and 2) malicious agents. Define the cooperative classification set $Y = \{Y_1, Y_2\}$. Agent $i \in Y_1$, while reducing its own energy consumption, also hopes that all "colleagues," i.e., its neighbors, can reduce their energy consumption, so it will selflessly claim its true states information to its neighbors. Agent $i \in Y_2$ does not care about the loss of others, and has a certain sense of selfishness. It fears that its loss will increase due to the exposure of its privacy in the first round of belief update. Therefore, agents belonging to Y_2 choose to deceive other agents. In this case, how one agent determines whether a neighbor is friendly or not is an important issue.

Reasonably, we assume that each agent i does not believe the information $m_{-i}(t, t+T)$ claimed by all its neighbors in the first round of belief update in order to reduce its own unnecessary loss, but it will record the information.

For each agent i, in the second round of belief update, the information $m_{-i}(t, t+T)$ received in the last round of belief update will be compared with the neighbors' states $x_j(t+T)$ observed at the previous moment t, and then choose whether to use the information $m_{-i}(t+T, t+2T)$ claimed by the neighbors in this round of belief update. The selection principle is as follows.

Case 1: $m_j(t, t+T) = x_j(t+T)$. Then, starting from the second round of belief update, agent i believes the information $m_j(t+(k-1)T, t+kT), k = 2, 3, 4 \ldots$, from neighbor j and takes it as the state of neighbor j, $x_j(t+kT)$, at time $t+kT$, $k = 2, 3, 4 \ldots$

Case 2: $m_j(t, t+T) \neq x_j(t+T)$. Then, starting from the second round of belief update, agent i still does not believe the information claimed by neighbor j. In this case we define the same normal distribution $N(\mu_{ij}^{\theta}, \mathrm{Cov}_{ij}^{\theta})$ as [11] for the state $x_j(t+T)$. The mean value of the normal distribution is taken as the estimate of j's state at the kth round of belief update: $\mu_{ij}^{\theta} = \hat{x}_j^{\theta}(t+kT)$, where

$$\hat{x}_j^{\theta}(t+kT) = e^{A(t+kT)}x_j(t) + \int_t^{t+kT} e^{-A(\tau-t-kT)}B\hat{u}_j^{i,\theta}(\tau)d\tau.$$

$\mathrm{Cov}_{ij}^{\theta}$ indicates the degree of the lack of confidence.

In addition to the above action, agent i will also stop transmitting true information to agent j from the moment it finds out itself is cheated. In this way, in the signal game of belief update, there is no unilateral cooperative game, i.e., once one party breaks the cooperation, the other party will do the same accordingly. In order to distinguish the agents who break cooperation actively and the agents who break cooperation passively, define $Z = \{Z_1, Z_2\}$, for $i \in Z_1$, agent i takes the initiative to break cooperation. For $i \in Z_2$, agent i is the passive one.

However, the choice between believing and deceiving should not be static in the actual game. The neighbor of agent i may shift from deception to cooperation, or from cooperation to deception due to problems such as glitches. Inspired by the Tit-for-Tat strategy in game theory [25], [26], we establish an information tradeoff principle to deal with the situation when a neighbor changes from cooperation to deception or from deception to cooperation.

1) Principle 1: The principle for dealing with the situation when a neighbor changes from cooperation to deception.

As long as $m_j(t+(k-1)T, t+kT) \neq x_j(t+kT)$ occurs during the k round of belief update, agent i will not trust the information from neighbor j and will instead use its own normal estimate of j until j satisfies Principle 2.

2) Principle 2: The principle for dealing with the situation when a neighbor changes from deception to cooperation.

In this situation, the neighbor may be the one who has been cheating i since the first round of belief update or the neighbor who once cooperated with i. Next, we discuss them separately.

1) For the neighbor who has been cheating i since the first round of belief update, i.e., for the neighbor j whose $m_j(t, t+T) \neq x_j(t+T)$ in the first round of belief update, agent i will ignore the information sent by j for a period of time, and use its own normal estimate value $\mu_{ij}^{\theta} = \hat{x}_j^{\theta}(t+kT), k = 2, 3, 4, \ldots$, as the state estimate of j at the kth round of belief update. However, agent i will record the information sent by j every time and make the continuous comparison.

If: $m_j(t+(k-1)T, t+kT) \neq x_j(t+kT), k = 1, 2, \ldots p$, $m_j(t+(k-1)T, t+kT) = x_j(t+kT), k = p+1, \ldots 2p$, then agent i will choose to trust the information sent by neighbor j at the $2p+1$st round of belief update, that is, agent i believes that $m_j(t+(k-1)T, t+kT) = x_j(t+kT), k = 2p+1$ holds.

Remark 6: If in the first round of belief update, agent i chooses cooperation, i.e., $i \in Y_1$, then from the second round of belief update, the relationship between other agents and i can be divided into the cooperative relationship and noncooperative relationship. Define the set of agents having cooperative relationship with i in the kth round of belief update be Y_{ik}^1, and the set of agents having noncooperative relationship with i as Y_{ik}^2. If $j \in Y_{ik}^1$, then

$$p\begin{pmatrix} x_i(t+kT)|x_{-j}(t+(k-1)T), \\ \theta, m_i(t+(k-1)T, t+kT) \end{pmatrix} = 1.$$

If $q \in Y_{ik}^2$, then $p(\begin{smallmatrix} x_i(t+kT)|x_{-i}(t+(k-1)T), \\ \theta, m_i(t+(k-1)T, t+kT) \end{smallmatrix})$ must be less than or equal to 1, where the equality is almost impossible to hold as it holds only when the expected value of the normal estimate is coincidently equal to the true value. So malicious agents will try to cooperate in order to find the actual type of its cost function.

2) For the neighbor agent j who once cooperated with i, suppose the last time j began establishing a cooperative relationship with i was in the gth round of update and it chose to deceive agent i from the qth round of belief update, i.e., $m_j(t+(k-1)T, t+kT) = x_j(t+kT), k = g, \ldots, q-1$, $m_j(t+(k-1)T, t+kT) \neq x_j(t+kT), k = q$.

However, after a period of time, it realizes that the effect of deception is no better than cooperation, so it starts to cooperate again. In this process, agent j will send true information to agent i. When the following conditions are satisfied: $m_j(t+(k-1)T, t+kT) \neq x_j(t+kT), k = q+1, \ldots, q+n$, $m_j(t+(k-1)T, t+kT) = x_j(t+kT), k = q+n+1, \ldots, q+3n$, agent i chooses to trust neighbor j again.

Remark 7: False information transmitted by neighbors that have never cooperated with i will never be adopted by i, so the belief update of i will not be affected by false information transmitted by these neighbors. However, when neighbors who had cooperated with i suddenly turned to deception, i will therefore use false information for one round of belief update. Therefore, when these neighbors want to cooperate with i again, they must cost more to gain trust than an agent that has never cooperated with i.

C. Extensive Game

Although the above-mentioned information tradeoff principle provides a way to deal with false information, it cannot avoid trauma when a neighbor suddenly switches from cooperation to deception. Therefore, it does not have the defensive ability or counterattack ability. We will show in this section that this problem can be solved when we let the information tradeoff principle become common knowledge in the graph.

In the first round of belief update, all agents agree to adopt the information tradeoff principle. Then all agents know the information tradeoff principle, and every agent knows that all agents know and adopt the information tradeoff principle. Then adopting information tradeoff principle becomes the common knowledge of the game.

When the friendly agents and the malicious agents make decisions, there will be two kinds of games that come from two situations. One situation is that two agents who have established a cooperative relationship make decisions, and

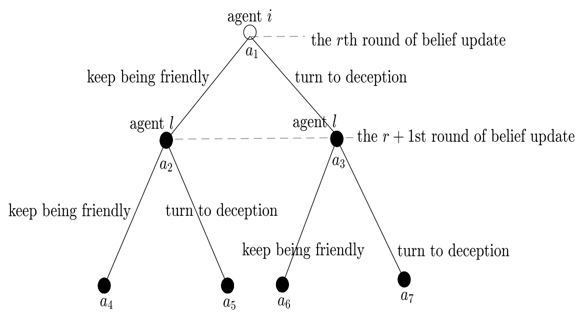

Fig. 2. Extensive game for the first situation.

the other is that two agents that do not have a cooperative relationship make decisions. In the following, we will analyze the games of the two situations and obtain the Nash equilibrium. The equilibrium decision of agents will show that the information tradeoff principle has sufficient defensive ability and counterattack ability.

Remark 8: It is worth noting that the extensive games discussed in this section are related to the game among agents during the process of belief update through communication interactions, rather than the graphical games mentioned earlier. Therefore, the Nash equilibria mentioned in this section refer to the Nash equilibrium decisions made by agents regarding whether to transmit true information during communication interaction games. In contrast, the Nash equilibria mentioned in Sections I and IV refer to the Nash equilibrium solutions for graphical games, i.e., the distributed optimal Nash controllers for the optimal consensus control problem.

The process of belief update is a process in which an agent needs to perform multiple rounds of interaction and cannot know in advance when the interaction will end. Therefore, the process of belief update can be regarded as a long-term game [24] between agents. In long-term game, there is always a chance of retaliation, such as no longer cooperating. Hence, malicious agents will suffer heavy losses in the long run.

1) Extensive Game for the First Situation: We build an extensive game between agent i and agent l who have established a cooperative relationship. In the kth round of belief update, when the friendly agent i chooses to remain friendly or turn to deception, it will form the game problem with l as shown in Fig. 2.

Point a_1 represented by a hollow circle at the top represents the beginning of the game. The "agent i" is marked on point a_1, indicating that the game begins with the choice of agent i in the rth round of belief update. The choices of agent i include "stay friendly to l" and "turn to cheat l," which are represented by directed branches labeled "keep being friendly" and "turn to deception," respectively. If agent i chooses to keep being friendly, the game goes from point a_1 to point a_2. Otherwise, the game goes from point a_1 to point a_3.

Points a_2 and a_3 are marked with "agent l," which means that agent l will make a decision when the game reaches point a_2 or a_3. That is, after i makes a decision, l will make a decision in the $r+1$st round of belief update. Agent l has two strategies, "stay friendly" and "turn to deception", represented by directed branches labeled keep being friendly and turn to deception, respectively. After agent l makes the corresponding decision, the game will reach the end point and obtain the corresponding income vector (c_1, c_2), where c_1 and c_2 are the incomes of i and l, respectively.

Let the income of adopting true information from one neighbor in one round of belief update be π while the income of adopting false information from one neighbor be $-\pi$. Let the income of using the normal estimation without adopting information be $\upsilon\pi, 0 < \upsilon < 1$.

Let the discount factor of the long-time game be γ. In game theory, the discount factor represents the degree of patience of a participant and $0 < \gamma < 1$ [24]. A larger γ means a higher probability that the participant will continue to play the game. If γ is equal to 0, it means that the participant is completely impatient. Moreover, the discount factor is the rate at which future earnings are converted to present value.

1) If agent i remains friendly to l in the rth round of belief update and agent l chooses to keep being friendly in the $r+1$st round of belief update, then agent l gets true information in the rth round of belief update and the income of l is π. In the $r+1$st round of belief update, both i and l get true information, and the incomes of i and l is π, respectively. Since the process of belief update can be regarded as a long-term game, the income of the $r+2$nd round of belief update and subsequent updates is $\gamma\pi + \gamma^2\pi + \gamma^3\pi + \cdots = \sum_{i=1}^{\infty}\gamma^i\pi$. So the income vector corresponding to point a_4 is $(\pi + ([\gamma\pi]/[1-\gamma]), 2\pi + [\gamma\pi/(1-\gamma)])$.
2) If agent i remains friendly to l in the rth round of belief update but agent l chooses to turn to deception in the $r+1$st round of belief update, then agent l gets true information in the rth round of belief update and the income of l is π. In the $r+1$st round of belief update, i adopts false information so the income is $-\pi$, while l gets true information so the income is π. Starting from the $r+2$nd round of belief update, the two agents no longer cooperate and the incomes are both $\gamma\upsilon\pi + \gamma^2\upsilon\pi + \gamma^3\upsilon\pi + \cdots = \sum_{i=1}^{\infty}\gamma^i\upsilon\pi$. So the income vector corresponding to point a_5 is $(-\pi + ([\upsilon\gamma\pi]/[1-\gamma]), 2\pi + [\upsilon\gamma\pi/(1-\gamma)])$.
3) If agent i turns to deceive l in the rth round of belief update and agent l still chooses to keep being friendly in the $r+1$st of belief update, then in the rth round of belief update agent l adopts false information and the income of l is $-\pi$. In the $r+1$st round of belief update, i gets true information and the income is π, while l adopts false information and the income is $-\pi$. In the $r+2$nd round of belief update and subsequent belief updates, the income of i is $\gamma\pi + \gamma^2\pi + \gamma^3\pi + \cdots = \sum_{i=1}^{\infty}\gamma^i\pi$ and the income of l is $\gamma\upsilon\pi + \gamma^2\upsilon\pi + \gamma^3\upsilon\pi + \cdots = \sum_{i=1}^{\infty}\gamma^i\upsilon\pi$. So the income vector corresponding to point a_6 is $(\pi + (\gamma\pi/[1-\gamma]), -2\pi + (\upsilon\gamma\pi/[1-\gamma]))$. It is worth noting that in this scenario, the behavior of agent j contradicts the information tradeoff principle, so in reality, j would not make such a decision. In the subsequent discussion, we demonstrate through the calculation of the incomes for different decisions that

TABLE I
STRATEGIC FORM GAME

strategic description	i choose to be friendly	i choose to be deceitful
{friendly, friendly}	$(\pi+\frac{\gamma\pi}{1-\gamma}, 2\pi+\frac{\gamma\pi}{1-\gamma})$	$(\pi+\frac{\gamma\pi}{1-\gamma}, -2\pi+\frac{\upsilon\gamma\pi}{1-\gamma})$
{friendly, deceitful}	$(\pi+\frac{\gamma\pi}{1-\gamma}, 2\pi+\frac{\gamma\pi}{1-\gamma})$	$(-\pi+\frac{\upsilon\gamma\pi}{1-\gamma}, -\pi+\upsilon\pi+\frac{\upsilon\gamma\pi}{1-\gamma})$
{deceitful, friendly}	$(-\pi+\frac{\upsilon\gamma\pi}{1-\gamma}, 2\pi+\frac{\upsilon\gamma\pi}{1-\gamma})$	$(\pi+\frac{\gamma\pi}{1-\gamma}, -2\pi+\frac{\upsilon\gamma\pi}{1-\gamma})$
{deceitful, deceitful}	$(-\pi+\frac{\upsilon\gamma\pi}{1-\gamma}, 2\pi+\frac{\upsilon\gamma\pi}{1-\gamma})$	$(-\pi+\frac{\upsilon\gamma\pi}{1-\gamma}, -\pi+\upsilon\pi+\frac{\upsilon\gamma\pi}{1-\gamma})$

this choice does not constitute a Nash equilibrium decision and would not be selected by agent j.

4) If agent i turns to deceive l in the rth round of belief update and agent l chooses to turn to deception in the $r+1$st round of belief update, then in the rth round of belief update l adopts false information and the income is $-\pi$. In the $r+1$st round of belief update, the agent i adopts false information so the income is $-\pi$, while the agent l discards false information so the income is $\upsilon\pi$. In the $r+2$nd round of belief update and subsequent belief updates, the incomes of i and l are both $\gamma\upsilon\pi+\gamma^2\upsilon\pi+\gamma^3\upsilon\pi+\cdots=\sum_{i=1}^{\infty}\gamma^i\upsilon\pi$. So the income vector corresponding to point a_7 is $(-\pi+([\upsilon\gamma\pi]/[1-\gamma]), -\pi+\upsilon\pi+[\upsilon\gamma\pi/(1-\gamma)])$.

In order to solve the Nash equilibrium corresponding to the game tree in Fig. 2, we transform the above extensive game into a strategic form game. We denote the strategy of agent l as $\{l_1, l_2\}$, where l_1 represents l's decision on node a_2 and l_2 represents l's decision on node a_3. Taking {friendly, friendly} as an example, it means that l remains friendly when i remains friendly and l remains friendly when i turns to deception. Abbreviate keep being friendly as friendly and turn to deception as deceitful. Then, a strategic description of the extensive game can be obtained as shown in Table I.

There are multiple Nash equilibria in this extensive game, but a reasonable Nash equilibrium is the subgame-perfect Nash equilibrium. For an extensive game, the subgame-perfect Nash equilibrium is generally used as the solution of the game. From Table I, three Nash equilibrium solutions {friendly, {friendly, friendly}}, {friendly, {friendly, deceitful}}, {deceitful, {deceitful, deceitful}} are obtained by the strategic form solution concept. Next, we adopt the subgame-perfect Nash equilibrium method to perform Nash refinement. The subgames include the game $\Gamma(a_2)$ with a_2 as the starting point and the game $\Gamma(a_3)$ with a_3 as the starting point. The Nash equilibrium of subgame $\Gamma(a_2)$ is that agent l chooses to keep being friendly and the Nash equilibrium of subgame $\Gamma(a_3)$ is that agent l chooses to turn to deception. Obviously, the Nash equilibrium of the original game {friendly, {friendly, deceitful}} constitutes a Nash equilibrium for all subgames, including the original game, so {friendly, {friendly, deceitful}} is a subgame-perfect Nash equilibrium. Since {friendly, {friendly, friendly}} and {deceitful, {deceitful, deceitful}} do not give Nash equilibria for the subgame $\Gamma(a_3)$ and $\Gamma(a_2)$, respectively, they are Nash equilibria but not subgame-perfect Nash equilibria. Therefore, {friendly, {friendly, deceitful}} is the Nash equilibria of this game.

2) Extensive Game for the Second Situation: We will prove in this game that a malicious agent would not remain deceitful all the time, and a friendly agent will cooperate with a malicious agent after the malicious agent confirms the information tradeoff principle.

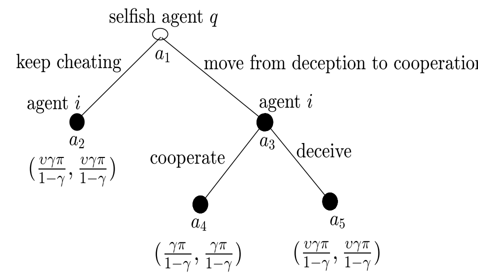

Fig. 3. Extensive game for the second situation.

We build an extensive game between agent i and agent q who do not have a cooperative relationship. If in the first round of belief update the agent q chooses deception, i.e., $q \in Y_2$, then starting from the second round of belief update, all its neighbors will no longer trust it and do not collaborate with it, i.e., cheat it.

When the malicious agent changes from being deceitful to being friendly, the neighbor agent has two choices. One is cooperation and the other is noncooperation. For this situation, the extensive game is shown in Fig. 3. The future income obtained when two agents cooperate is recorded as γ_2, and the income vector is $([\gamma\pi]/[1-\gamma], [\gamma\pi]/[1-\gamma])$. When not cooperating, the income vector is $([\upsilon\gamma\pi]/[1-\gamma], [\upsilon\gamma\pi]/[1-\gamma])$. In this game, (move from deception to cooperation, cooperate) is the dominant strategy and thus is the Nash equilibrium of the game.

Remark 9: We can draw the following conclusion from the two games of these two situations. Once the information tradeoff principle becomes common knowledge, the friendly agents will not actively turn to deception. A malicious agent will not remain deceitful all the time, and a friendly agent will cooperate with the malicious agent if the malicious agent satisfies Principle 2 of the information tradeoff principle.

D. Convergence Analysis

In this section, we will prove theoretically that the proposed update method has a faster convergence rate than the Bayesian belief update method [11] when malicious neighbors exist.

Theorem 2: For any positive integer $k \geq 1$, any states information x_{-i} and m_{-i}, when adopting the information tradeoff principle becomes the common knowledge of the game

$$p\begin{pmatrix}\theta | x_{-i}(t+kT), x_{-i}(t+(k-1)T), \\ \theta_i, m_{-i}(t+(k-1)T, t+kT)\end{pmatrix} \geq p(\theta | x_{-i}(t+kT), x_{-i}(t+(k-1)T), \theta_i)$$

holds for all θ.

Proof: Based on the analysis in Section III-C, we conclude that all agents have given long-term and objective consideration to their own interests. Finally, all the information transmitted in the communication graph will be true. Denote such a moment as T_λ and let $T_i, i = 1, 2, 3, \ldots$, be the time instants when all i's neighbors transmit true information. Then $T_\lambda = \sup\{T_i\}$.

So, for the belief update before time T_λ, if j belongs to Y_{ik}^1, then i believes the information from j. According to Remarks 6 and 9, the information transmitted by j is the true information. Then

$$p\begin{pmatrix} x_j(t+kT)|x_{-i}(t+(k-1)T), \\ \theta, m_j(t+(k-1)T, t+kT) \end{pmatrix} = 1$$
$$\geq p(x_j(t+kT)|x_{-i}(t+(k-1)T), \theta).$$

If j belongs to Y_{ik}^2, i does not believe the information from j. Then

$$p\begin{pmatrix} x_j(t+kT)|x_{-i}(t+(k-1)T), \\ \theta, m_j(t+(k-1)T, t+kT) \end{pmatrix}$$
$$= p(x_j(t+kT)|x_{-i}(t+(k-1)T), \theta).$$

Therefore

$$\begin{aligned} &p\begin{pmatrix} \theta|x_{-i}(t+kT), x_{-i}(t+(k-1)T), \\ \theta_i, m_{-i}(t+(k-1)T, t+kT) \end{pmatrix} \\ &= \frac{\prod_{j\in Y_{ik}^1} p\begin{pmatrix} x_j(t+kT)|x_{-i}(t+(k-1)T), \\ \theta, m_j(t+(k-1)T, t+kT) \end{pmatrix}}{p(x_{-i}(t+kT)|x_{-i}(t+(k-1)T), \theta_i)} \\ &\times p(\theta|x_{-i}(t+(k-1)T), \theta_i) \\ &\times \prod_{j\in Y_{ik}^2} p\begin{pmatrix} x_j(t+kT)|x_{-i}(t+(k-1)T), \\ \theta, m_j(t+(k-1)T, t+kT) \end{pmatrix} \\ &\geq \frac{\prod_{j\in Y_{ik}^1} p\big((x_j(t+kT)|x_{-i}(t+(k-1)T), \theta)\big)}{p(x_{-i}(t+kT)|x_{-i}(t+(k-1)T), \theta_i)} \\ &\times p(\theta|x_{-i}(t+(k-1)T), \theta_i) \\ &\times \prod_{j\in Y_{ik}^2} p(x_j(t+kT)|x_{-i}(t+(k-1)T), \theta) \\ &= p(\theta|x_{-i}(t+kT), x_{-i}(t+(k-1)T), \theta_i) \end{aligned}$$

holds for all θ.

For the belief update after time T_λ, all neighbors j belong to Y_{ik}^1. Then for any neighbor j

$$p\begin{pmatrix} x_j(t+kT)\ |x_{-i}(t+(k-1)T), \\ \theta, m_j(t+(k-1)T, t+kT) \end{pmatrix} = 1$$
$$\geq p\big(x_j(t+kT)\ |x_{-i}(t+(k-1)T), \theta\big).$$

In the same way, we can prove

$$p\begin{pmatrix} \theta|x_{-i}(t+kT), x_{-i}(t+(k-1)T), \\ \theta_i, m_{-i}(t+(k-1)T, t+kT) \end{pmatrix}$$
$$\geq p(\theta|x_{-i}(t+kT), x_{-i}(t+(k-1)T), \theta_i)$$

holds for all θ.

In summary, for any positive integer $k \geq 1$

$$p\begin{pmatrix} \theta|x_{-i}(t+kT), x_{-i}(t+(k-1)T), \\ \theta_i, m_{-i}(t+(k-1)T, t+kT) \end{pmatrix}$$
$$\geq p(\theta|x_{-i}(t+kT), x_{-i}(t+(k-1)T), \theta_i)$$

holds for all θ. Then we complete the proof.

From Theorem 2 we can see that with the same initial condition

$$p\begin{pmatrix} \theta^*|x_{-i}(t+kT), x_{-i}(t+(k-1)T), \\ \theta_i, m_{-i}(t+(k-1)T, t+kT) \end{pmatrix}$$
$$\geq p\big(\theta^*|x_{-i}(t+kT), x_{-i}(t+(k-1)T), \theta_i\big)$$

holds for all positive integers $k \geq 1$, where the equality holds if and only if the expected values of the normal estimates are coincidently equal to the true values for all k. Therefore, the belief update method with the information tradeoff principle will have a faster convergence rate to the true type θ^*.

In the consensus problem for frequency and voltage restoration of microgrids, the different degrees of actuation/propulsion faults and parameter disturbances will affect the control effort required for each node to achieve consensus. However, due to the uncertainty associated with false injection attacks in communication between grid nodes and the varying degrees of faults/parameter perturbations, nodes are unable to design the correct optimal controller to achieve the goal with minimal energy consumption. In such scenarios, the method proposed in this article, compared to traditional approaches, enables the system to design the correct optimal controller more quickly and achieve the objective with reduced energy consumption.

E. Improved Trust Mechanism

In this section, we further speed up the belief update convergence by speeding up the agent's judgment on an agent that switches from deception to cooperation. In the previous section, before friendly agents cooperate with the repentant malicious neighbor, the malicious neighbor may have already delivered lots of true information during the cooperative tentative phase. However, the friendly agents do not make full use of this information. In this section, we design a new mechanism that enables the friendly agent to use the information before malicious neighbors satisfy Principle 2 of the information tradeoff principle. Suppose that the neighbors that have formed a cooperative relationship with agent i in the kth round of belief update include $l_1^i(k), l_2^i(k), \ldots, l_{m_k}^i(k)$. Let the number of consecutive cooperation between the above neighbor and agent i be $r(l_1^i(k)), r(l_2^i(k)), \ldots, r(l_{m_k}^i(k))$, respectively. Let the number of times each neighbors have fooled agent i be $c(l_1^i(k)), c(l_2^i(k)), \ldots, c(l_{m_k}^i(k))$. Suppose the neighbor $l_j^i(k)$ deceived i in d_1^jth, d_2^jth, $\ldots, d_p^j$th, $\ldots, d_{c(l_{m_k}^i(k))}^j$th round of belief update, respectively. Design a trust index $\lambda_{ij} = r(l_j^i(k)) + \sum_{p=1}^{c(l_j^i(k))} (-1/[k-d_p^j])$.

We then find out the most trusted neighbors of agent i by finding the neighbors who make λ_{ij} largest. If g is one of the most trusted neighbors of i and g judges that a malicious agent l has turned from deceitful to friendly, that is, l satisfies g's information tradeoff principle. Then the agent i will start to adopt the information transmitted by the malicious agent l before l satisfies the information tradeoff principle of i. Thereby, the convergence rate of belief update can be faster.

IV. Stability and Distributed Global Bayesian Nash Equilibrium

Remark 10: In the process of belief update, the graph topology of the system may switch due to the tradeoff of information. But since the equilibrium of the extensive games in Section III-C proves that all agents will cooperate with each other in the end, the final graph topology will be the same as the initial topology and be fixed.

Next we show that when all agents cooperate with each other, the global error dynamics are asymptotically stable and the optimal strategies of all agents will form a Nash equilibrium.

Let

$$\delta = \left[\delta_1^{\mathrm{T}}, \ldots, \delta_N^{\mathrm{T}}\right]^{\mathrm{T}}. \tag{20}$$

According to Remark 10, the final graph topology will be the same as the initial topology. Substituting (12) into $\dot{\delta}_i$ yields

$$\begin{aligned}\dot{\delta}_i &= \Big(A - (d_i+g_i)^2 B(\mathbb{E}R_{ii})^{-1}(B^{\mathrm{T}}\mathbb{E}\tilde{P}_i)\Big)\delta_i \\ &\quad + \sum_{j\in N_i} a_{ij}(d_j+g_j)B(\mathbb{E}R_{jj})^{-1}(B^{\mathrm{T}}\mathbb{E}\tilde{P}_j)\delta_j. \end{aligned}\tag{21}$$

Then the global error dynamics is

$$\dot{\delta} = \Big((I\otimes A) - ((L+G)\otimes B)\ \bar{R}^{-1}\big((D+G)\otimes B^T\big)\tilde{P}\Big)\delta \tag{22}$$

where $\bar{R} = \mathrm{diag}_i\{\mathbb{E}R_{ii}\}$ and $\tilde{P} = \mathrm{diag}_i\{\mathbb{E}\tilde{P}_i\}$.

Similar to [12, Assumption 3], the following Assumption is given to analyze the consensus.

Assumption 3 [12, p. 7]: The matrices R_{ii}^{θ} in (6) are selected such that $\mathbb{E}R_{ii} = fR$ for all agents i, where the constant $f > 0$ and the matrix $R > 0$.

Theorem 3: When Assumptions 1-3 hold, the global error system δ in (22) is asymptotically stable, i.e., all agents achieve consensus by the optimal strategy in (12).

Proof: Please see the Appendix-C.

Lemma 2: Given the expected Hamiltonian function (7) and the optimal strategy (12), then

$$\begin{aligned}\mathbb{E}H_i(\delta_i, u_i, u_{-i}, \theta) &= \mathbb{E}H_i\big(\delta_i, u_i^*, u_{-i}, \theta\big) \\ &\quad + \sum_{\theta\in\Theta} p(\theta \mid \theta_i, m_{-i}) \sum_{j\in\mathcal{N}_i} \big(u_i - u_i^*\big)^T R_{ii}^{\theta}\big(u_i - u_i^*\big).\end{aligned}\tag{23}$$

Proof: Please see the Appendix-D.

Theorem 4: Suppose Assumptions 1 and 2 hold. Let $\tilde{V}_i^{\theta}$ take the form (11) and solve BHJ equation (10). Then the control strategies $(u_1^*, u_2^*, \ldots, u_N^*)$ provide a global Nash equilibrium after all malicious agents transmitting true information.

Proof: Please see the Appendix-E.

Remark 11: The Nash Equilibrium of the Bayesian graphical game has been given in [11]. From the simulations given in [11], the Nash Equilibrium in [11] is not distributed because it depends on the state information of their neighbors' neighbors. However, the Nash Equilibrium in this article is distributed because there exist distributed value functions $\tilde{V}_i^{\theta}(\delta_i)$ that guarantee the solvability of BHJ equations.

V. Simulation and Comparative Analysis

In this section, we conduct the simulation to show the effectiveness of the belief update method with the information tradeoff principle and compare it with traditional Bayesian belief update method.

A. Setup of the Simulation

Consider an unmanned aerial vehicle (UAV) swarm consisting of one virtual leader and five quad-rotor UAVs connected by the communication graph as shown in Fig. 4. The virtual leader is not a real UAV but represents a goal point given to the UAV swarm. Only UAV 1 can receive the goal point x_0.

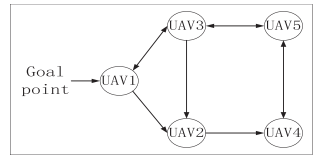

Fig. 4. Topology of the UAVs.

Referencing [27], let $x_i = [x_{ip}, \dot{x}_{ip}, \psi]^T$, $y_i = [y_{i1}, \dot{y}_{i1}, \phi]^T$, where x_{ip} and y_{ip} are positions in the xy-plane in the inertial frame, $\dot{x}_{ip}$ and $\dot{y}_{ip}$ are velocities, ψ and ϕ are rotation angles. Then the simplified quad-rotor models for the x-direction and y-direction are given as $\dot{x}_i = A_x x_i + B_x u_{ix}$, $\dot{y}_i = A_y y_i + B_y u_{iy}$, $i = 1, 2, 3, 4, 5$, where $A_x = [0\ 1\ 0; 0\ 0\ 9.8; 0\ 0\ -5.018]$, $B_x = [0\ 0\ 0; 0\ 0\ 0; 0\ 0\ 5.233]$, $A_y = [0\ 1\ 0; 0\ 0\ -9.8; 0\ 0\ -5.083]$, $B_y = [0\ 0\ 0; 0\ 0\ 0; 0\ 0\ 5.516]$, u_{ix} and u_{iy} are the desired inputs to the on-board controller. As pointed out in [27], these linear models are only valid when the rotation angles are small (typically in the range of $\pm 10°$).

Let $x_0 = [0, 0, 0]^T$, $y_0 = [0, 0, 0]^T$, then the simulation can then be viewed as the process of the UAV swarm flying in formation to a destination.

Considering the presence of false data injection attacks in the communication topology, each UAV hopes to achieve formation tracking while minimizing fuel consumption in uncertain weather or terrain environments. Then this formation tracking control problem of UAVs systems can be addressed in the context of a MAS Bayesian graphical game.

The UAVs systems can be regarded as a MAS with UAVs being agents indexed by $1, \ldots, N$. All the agents have a certain type except Agent 1. Agent 1 has a 0.1 probability of type 1 θ^1 and a 0.9 probability of type 2 θ^2. Agent 1, Agent 2, Agent 3, and Agent 4 $\in Y_1$, while Agent 5 $\in Y_2$. $a_{ij} = 1 \quad \forall i, j = 1, \ldots, N$. Two cost functions in the form of (6) were established, respectively, in the x-direction and the y-direction to solve the optimal controller. The parameter matrices of the two cost functions are the same and as follows: $Q_{ii}^{\theta^1} = [10\ 0\ 0; 0\ 1\ 0; 0\ 0\ 0.1]$, $Q_{ii}^{\theta^2} = [0.1\ 0\ 0; 0\ 0.01\ 0; 0\ 0\ 0.001]$, $R_{ii}^{\theta^1} = R_{ii}^{\theta^2} = 4*10^8 I$, $R_{ij}^{\theta^1} = R_{ij}^{\theta^2} = 10^{15} I$. The above parameters can ensure that at no point in the simulation did any angle of the quadrotors exceed 9°, thus maintaining the small angle assumption of the linear model.

B. Comparative Simulation of Belief Update

In the simulation, assume Agent 1 is in type 1 and the belief update period of the graph is 0.05 s. Agents that establish a cooperative relationship can use the true state information transmitted by the other. Otherwise, only the normal estimate value in Section III-B can be used. The belief update of agents is shown in Figs. 5–7. All the agents have a certain type except Agent 1. Agent 1 knows its own type. Then Agent 1 knows the actual global type. So we do not draw the belief update curve

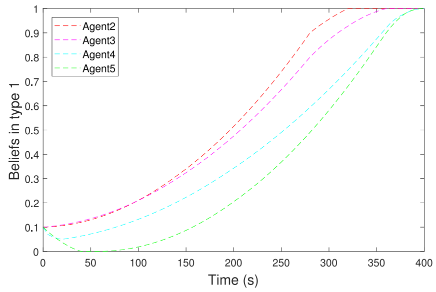

Fig. 5. Traditional Bayesian belief update method.

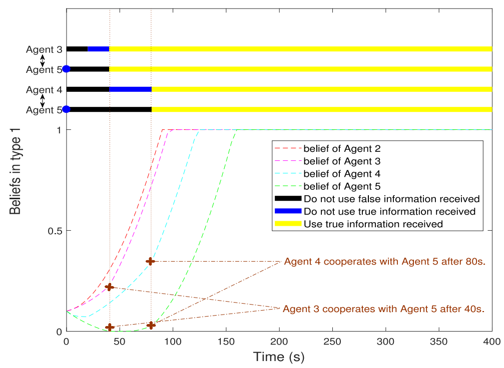

Fig. 6. Belief update method with information tradeoff principle.

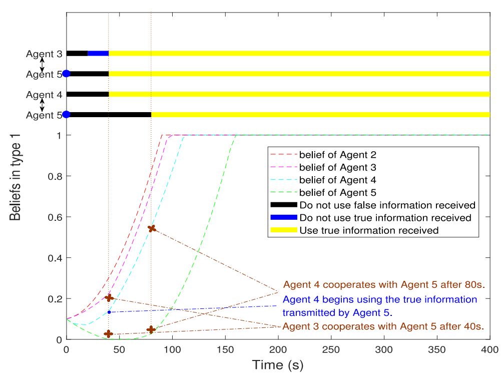

Fig. 7. Belief update method with the improved trust mechanism.

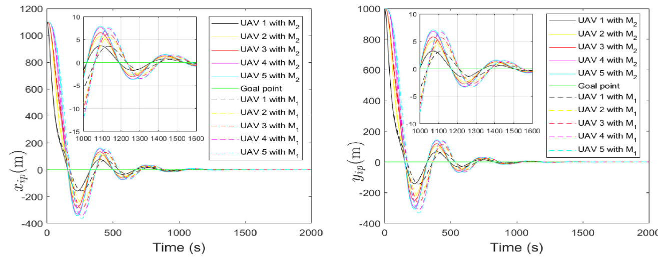

Fig. 8. Trajectories of UAVs.

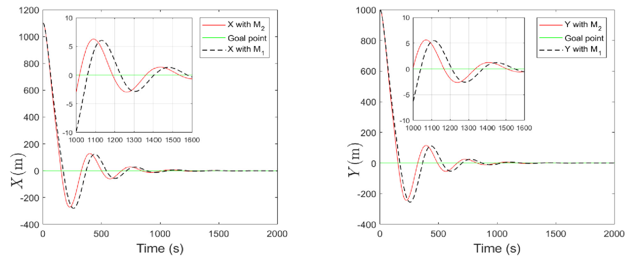

Fig. 9. Trajectories of X and Y.

of Agent 1. Suppose that Agent 5 transmits false information to Agent 3 and Agent 4 at 0 s. According to Remark 9, Agent 5 will not remain deceitful all the time. Suppose Agent 5 begins to transmit true information to Agent 3 at 20 s, and to Agent 4 at 40 s. Fig. 6 (Fig. 7) shows the information interactions between Agent 3 (Agent 4) and Agent 5 in detail. The four curves in Figs. 5–7 represent the evolution of the estimated probabilities of the four agents about θ^*. In Figs. 8 and 9, the Traditional Bayesian belief update method is abbreviated as M_1, the belief update method with information tradeoff principle is abbreviated as M_2, and the belief update method with the improved trust mechanism is abbreviated as M_3.

1) According to Remark 5, when using the traditional Bayesian belief update, agents will discard all the information from their neighbors without any judgment, and only use the neighbors' historical states for estimation. Then the agents cannot make use of the information from neighbors.
2) According to Remark 6, when using the belief update method with information tradeoff principle, Agent 2 can use true information from its neighbors after 0.05 s. Agent 3 can use true information from Agent 1 after 0.05 s. Agent 4 can use true information from Agent 2 after 0.05 s. Agent 3 will cooperate with Agent 5 after 40 s and then use true information from Agent 5. Agent 4 will cooperate with Agent 5 after 80 s and then use true information from Agent 5. Correspondingly, Agent 5 can use true information from Agent 3 after 40 s and true information from Agent 4 after 80 s. From Figs. 5 and 6, we can see that the belief of all agents converge to probability 1 before 200 s in Fig. 6 while all agents converge to probability 1 after 350 s in Fig. 5. Then all agents have a faster convergence rate than they use the traditional Bayesian belief update. This is consistent with Theorem 2.
3) When using the belief update method with the improved trust mechanism in Section III-E, Agent 4 will cooperate with Agent 5 after 80 s but use true information from Agent 5 after 40 s. Then Agent 4 will have a faster convergence rate than it uses the belief update method with information tradeoff principle. By comparing the two sky blue curves in Figs. 6 and 7, we can know that the improved trust mechanism enables Agent 4 to use the true information from Agent 5 earlier than the information tradeoff principle, which yields a faster convergence rate.

In Fig. 8, we compare the state trajectories of all UAVs when using the belief update method with the information tradeoff principle (the solid lines) and when using the Bayesian belief update method (the dashed lines). In order to compare the convergence rate in a clearer way, let $X = (x_1 + x_2 + x_3 + x_4 + x_5)/5$ and $Y = (y_1 + y_2 + y_3 + y_4 + y_5)/5$. The simulation stipulates that the UAV reaches the target point if it is within five meters of the destination. The trajectors of X and Y are shown in Fig. 9, from which we can see that the belief update method with information tradeoff principle enables UAVs to reach the goal point, i.e., achieve consensus at a faster rate.

VI. Conclusion

This article analyzes the optimal consensus problem where there exist some malicious agents. The distributed global Nash equilibrium of the MAS Bayesian graphical game is solved to achieve consensus. An information tradeoff principle and a belief update method are designed to solve the belief update problem in the presence of malicious neighbors. It is proved that this new belief update method has a faster convergence rate than the Bayesian belief update method when malicious neighbors exist.

In practical application, there will be situations in which the agent does not know its own type. For instance, in the previously mentioned consensus problem for frequency and voltage restoration of microgrids, node i may not be able to confirm immediately which specific hardware component or power line has malfunctioned when it experiences an actuation/propulsion fault. In Section II, we define the degree of faults and parameter perturbations for each node as the node's type. Therefore, when node i lacks a comprehensive understanding of the location of its own fault, it leads to the issue of node i being uncertain about its own type. The relevant research on this special case can be carried out in the future.

Appendix

A. Proof of Lemma 1

Substituting $\tilde{V}_i^\theta(\delta_i) = \delta_i^{\mathrm{T}} \tilde{P}_i^\theta \delta_i$ into (9), (12) is yielded. Substituting $\tilde{V}_i^\theta(\delta_i) = \delta_i^{\mathrm{T}} \tilde{P}_i^\theta \delta_i$ into (10), then (10) can be rewritten as

$$
\begin{aligned}
&\sum_{\theta\in\Theta} p(\theta \mid \theta_i, m_{-i})(\sum_{j\in\mathcal{N}_i} \delta_i^T(\tilde{P}_i^{\theta\mathrm{T}} A + A^T \tilde{P}_i^\theta + Q_i^\theta \\
&\quad - (d_i+g_i)^2 \tilde{P}_i^{\theta\mathrm{T}} B(\mathbb{E}R_{ii})^{-1} B^T \mathbb{E}\tilde{P}_i)\delta_i \\
&\quad + \sum_{j\in\mathcal{N}_i} \delta_i^T(\tilde{Q}_{ij}^\theta + N_i a_{ij}(d_j+g_j)\tilde{P}_i^{\theta\mathrm{T}} B(\mathbb{E}R_{jj})^{-1} B^T \mathbb{E}\tilde{P}_j)\delta_j \\
&\quad + \sum_{j\in\mathcal{N}_i} \delta_j^T(\tilde{Q}_{ij}^{\theta\mathrm{T}} + N_i a_{ij}(d_j+g_j)\tilde{P}_j^{\theta\mathrm{T}} B(\mathbb{E}R_{jj})^{-1} B^T \mathbb{E}\tilde{P}_i)\delta_i \\
&\quad + \sum_{j\in\mathcal{N}_i} \delta_j^T(\hat{Q}_{ij}^\theta - a_{ij}(d_j+g_j)^2 \tilde{P}_j^{\theta\mathrm{T}} B(\mathbb{E}R_{jj})^{-1}(\mathbb{E}R_{ij})(\mathbb{E}R_{jj})^{-1} \\
&\quad \cdot B^T \mathbb{E}\tilde{P}_j)\delta_j) = 0. \qquad (24)
\end{aligned}
$$

From [12], it can be known that there exists unique solution $\mathbb{E}\tilde{P}_i$ to (24). Equation (12) is unique due to the uniqueness of $\mathbb{E}\tilde{P}_i$.

B. Proof of Theorem 1

From [12], it can be known that there always exist a unique positive definite solution $\mathbb{E}\tilde{P}_i$ of (13) if $(A, \sqrt{\sum_{\theta\in\Theta} p(\theta \mid \theta_i, m_{-i}) Q_i^\theta})$ is observable and (A, B) is stabilizable. Substituting (13)–(15) into (24), the existence of the solution to the BHJ equation (10) is proved.

C. Proof of Theorem 3

The following two Lemmas from [12] are needed.

Lemma 3 [12, p. 7]: If Assumptions 2 and 3 hold, there exists a positive definite matrix $F = \mathrm{diag}_i\{f_i\}$ such that $S_1 := ((L+G)(D+G)F^{-1} + F^{-1}(D+G)^T(L+G)^T) \otimes BR^{-1}B^T \geq 0$.

Lemma 4 [12, p. 7]: If Assumptions 2 and 3 hold, choose f satisfying $(\sum_{\theta\in\Theta} p(\theta \mid \theta_i, m_{-i}) Q_i^\theta) - (1/f)(d_i+g_i)^2 \mathbb{E}\tilde{P}_i^{\mathrm{T}} BR^{-1}B^T \mathbb{E}\tilde{P}_i > 0$. Then $S_2 := -P_F^{-1}(I \otimes A^{\mathrm{T}}) - (I\otimes A)\, P_F^{-1} > 0$, where $P_F = (F \otimes I)\, \tilde{P}$.

Next, we give the proof of the Theorem.

Proof: According to Lemmas 3 and 4, we can obtain

$$
\begin{aligned}
&-\frac{1}{f}S_1 - S_2 < 0 \\
&\quad -\frac{1}{f}\left((L+G)(D+G)F^{-1} \otimes BR^{-1}B^{\mathrm{T}}\right) + (I\otimes A)P_F^{-1} < 0 \\
&\quad -\frac{1}{f}\left(F^{-1}(D+G)^{\mathrm{T}}(L+G)^{\mathrm{T}} \otimes BR^{-1}B^{\mathrm{T}}\right) + P_F^{-1}\left(I\otimes A^{\mathrm{T}}\right) \\
&\quad -\frac{1}{f}\tilde{P}^{-1}(F^{-1}\otimes I)\tilde{P}\left(((D+G)^T(L+G)^T) \otimes BR^{-1}B^{\mathrm{T}}\right) \\
&\quad + P_F^{-1}(I\otimes A) \\
&\quad -\frac{1}{f}\left(((L+G)(D+G)) \otimes BR^{-1}B^{\mathrm{T}}\right) + (I\otimes A)P_F^{-1} < 0 \\
&\quad P_F^{-1}\left((I\otimes A) - ((L+G) \otimes B)\bar{R}^{-1}\left((D+G) \otimes B^{\mathrm{T}}\right)\tilde{P}\right)^{\mathrm{T}} \\
&\quad + \left((I\otimes A) - ((L+G) \otimes B)\bar{R}^{-1}\left((D+G) \otimes B^{\mathrm{T}}\right)\tilde{P}\right)P_F^{-1} < 0 \\
&\quad \mathrm{Re}_{\max}\left((I\otimes A) - ((L+G) \otimes B)\bar{R}^{-1}\left((D+G) \otimes B^{\mathrm{T}}\right)\tilde{P}\right)^{\mathrm{T}} < 0.
\end{aligned}
$$

So the global error δ is asymptotically stable.

D. Proof of Lemma 2

$$
\begin{aligned}
&\mathbb{E}H_i(\delta_i, u_i, u_{-i}, \theta) = \sum_{\theta\in\Theta} p(\theta \mid \theta_i, m_{-i}) \\
&\quad \times \sum_{j\in\mathcal{N}_i} \begin{pmatrix} \frac{\partial \tilde{V}_i^{\theta\mathrm{T}}}{\partial \delta_i}\dot{\delta}_i + \delta_{ij}^{\mathrm{T}} Q_{ij}^\theta \delta_{ij} + u_i^{\mathrm{T}} R_{ii}^\theta u_i - a_{ij} u_j^{\mathrm{T}} R_{ij}^\theta u_j \\ + u_i^{*\mathrm{T}} R_{ii}^\theta u_i^* - u_i^{*\mathrm{T}} R_{ii}^\theta u_i^* + \frac{\partial \tilde{V}_i^{\theta\mathrm{T}}}{\partial \delta_i}(d_i+g_i)Bu_i^* \\ - \frac{\partial \tilde{V}_i^{\theta\mathrm{T}}}{\partial \delta_i}(d_i+g_i)Bu_i^* \end{pmatrix} \\
&\quad = \mathbb{E}H_i(\delta_i, u_i^*, u_{-i}, \theta) \\
&\qquad + \sum_{\theta\in\Theta} p(\theta \mid \theta_i, m_{-i}) \sum_{j\in\mathcal{N}_i} \left(\frac{\partial \tilde{V}_i^{\theta\mathrm{T}}}{\partial \delta_i}(d_i+g_i)Bu_i \right. \\
&\qquad \left. + u_i^{\mathrm{T}} R_{ii}^\theta u_i - u_i^{*\mathrm{T}} R_{ii}^\theta u_i^* - \frac{\partial \tilde{V}_i^{\theta\mathrm{T}}}{\partial \delta_i}(d_i+g_i)Bu_i^* \right).
\end{aligned}
$$

According to (9), the following equations hold:

$$\begin{aligned}&\sum_{\theta\in\Theta} p(\theta \mid \theta_i, m_{-i})(u_i-u_i^*)^T R_{ii}^\theta (u_i-u_i^*)\\&=\sum_{\theta\in\Theta} p(\theta \mid \theta_i, m_{-i})\left(u_i^{\mathrm{T}} R_{ii}^\theta u_i + u_i^{*\mathrm{T}} R_{ii}^\theta u_i^* - 2u_i^{*\mathrm{T}} R_{ii}^\theta u_i\right)\\&\sum_{\theta\in\Theta} p(\theta|\theta_i, m_{-i})(d_i+g_i)\frac{\partial \tilde{V}_i^{\theta\mathrm{T}}}{\partial\delta_i} B u_i^*\\&=\sum_{\theta\in\Theta} p(\theta|\theta_i, m_{-i})\left(-2u_i^{*\mathrm{T}} R_{ii}^\theta u_i^*\right)\mathbb{E}H_i(\delta_i, u_i, u_{-i}, \theta)\\&=\mathbb{E}H_i(\delta_i, u_i^*, u_{-i}, \theta)+\sum_{\theta\in\Theta} p(\theta \mid \theta_i, m_{-i})\\&\sum_{j\in\mathcal{N}_i}(u_i-u_i^*)^T R_{ii}^\theta (u_i-u_i^*).\end{aligned}$$

E. Proof of Theorem 4

According to Remark 10, there exists an instant T_t, after which all agents transmit true information to each other. Because the global error system is asymptotically stable, the local error $\lim_{t\to+\infty}\delta_i(t) = 0$. Then $\tilde{V}_i^\theta(\delta_i(\infty)) = 0$. Consequently

$$\begin{aligned}&\mathbb{E}J_i(\delta_i(T_t), \delta_{-i}(T_t), u_i, u_{-i})\\&=\sum_{\theta\in\Theta} p(\theta \mid \theta_i, m_{-i}) J_i^\theta(\delta_i(T_t), \delta_{-i}(T_t))\\&=\sum_{\theta\in\Theta}(p(\theta \mid \theta_i, m_{-i}))\\&\sum_{j\in\mathcal{N}_i}\bigg(\int_{T_t}^{\infty}\left(\delta_{ij}^{\mathrm{T}} Q_{ij}^\theta \delta_{ij} + u_i^{\mathrm{T}} R_{ii}^\theta u_i - a_{ij} u_j^{\mathrm{T}} R_{ij}^\theta u_j\right)dt\\&\quad+\tilde{V}_i^\theta(\delta_i(T_t)) - \tilde{V}_i^\theta(\delta_i(T_t)) + \tilde{V}_i^\theta(\delta_i(\infty))\bigg)\\&=\int_{T_t}^{\infty}\mathbb{E}H_i(\delta_i, \delta_{-i}, u_i, u_{-i})dt\\&\quad+\sum_{\theta\in\Theta} p(\theta \mid \theta_i, m_{-i})\sum_{j\in\mathcal{N}_i}\tilde{V}_i^\theta(\delta_i(T_t))\\&=\int_{T_t}^{\infty}\mathbb{E}H_i\big(\delta_i, \delta_{-i}, u_i^*, u_{-i}\big)dt\\&\quad+\sum_{\theta\in\Theta} p(\theta \mid \theta_i, m_{-i})\sum_{j\in\mathcal{N}_i}\tilde{V}_i^\theta(\delta_i(T_t))\\&\quad+\sum_{\theta\in\Theta}(p(\theta \mid \theta_i, m_{-i}))\sum_{j\in\mathcal{N}_i}\int_{T_t}^{\infty}(u_i-u_i^*)^T R_{ii}^\theta(u_i-u_i^*)dt.\end{aligned}$$

Obviously $\mathbb{E}J_i(\delta_i, \delta_{-i}, u_i^*, u_{-i}) \le \mathbb{E}J_i(\delta_i, \delta_{-i}, u_i, u_{-i})$. If $u_{-i} = u_{-i}^*$, then $\mathbb{E}J_i(\delta_i, \delta_{-i}, u_i^*, u_{-i}^*) \le \mathbb{E}J_i(\delta_i, \delta_{-i}, u_i, u_{-i}^*)$ holds for $\forall u_i, i = 1, 2, \ldots, N$. It is proved that after all malicious agents transmitting true information, the global Nash equilibrium can be achieved.

References

[1] S. Ben Meskina, N. Doggaz, M. Khalgui, and Z. Li, "Multiagent framework for smart grids recovery," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 47, no. 7, pp. 1284–1300, Jul. 2017.

[2] B. Chen, J. Hu, Y. Zhao, and B. K. Ghosh, "Finite-time velocity-free rendezvous control of multiple AUV systems with intermittent communication," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 52, no. 10, pp. 6618–6629, Oct. 2022.

[3] Y.-W. Wang, Y.-W. Wei, X.-K. Liu, N. Zhou, and C. G. Cassandras, "Optimal persistent monitoring using second-order agents with physical constraints," *IEEE Trans. Autom. Control*, vol. 64, no. 8, pp. 3239–3252, Aug. 2019.

[4] X. Yao, Y. Liu, and W. Zhao, "Adaptive boundary vibration control and angle tracking consensus of networked flexible Timoshenko manipulator systems," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 53, no. 5, pp. 2949–2960, May 2023.

[5] X.-Z. Jin, W.-W. Che, Z.-G. Wu, and C. Deng, "Robust adaptive general formation control of a class of networked quadrotor aircraft," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 52, no. 12, pp. 7714–7726, Dec. 2022.

[6] X. Cai, F. Xiao, and B. Wei, "Resilient Nash equilibrium seeking in multiagent games under false data injection attacks," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 53, no. 1, pp. 275–284, Jan. 2023.

[7] W. Ben-Ameur, P. Bianchi, and J. Jakubowicz, "A total variation based approach for robust consensus in distributed networks," in *Proc. 52nd IEEE Conf. Decision Control*, 2013, pp. 1690–1695.

[8] H. Sun, Y. Liu, and F. Li, "Optimal consensus via distributed protocol for second-order multiagent systems," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 51, no. 10, pp. 6218–6228, Oct. 2021.

[9] J. Shi, D. Yue, and X. Xie, "Optimal leader–follower consensus for constrained-input multiagent systems with completely unknown dynamics," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 52, no. 2, pp. 1182–1191, Feb. 2022.

[10] C. An, H. Su, and S. Chen, "Inverse-optimal consensus control of fractional-order multiagent systems," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 52, no. 8, pp. 5320–5331, Aug. 2022.

[11] V. G. Lopez, Y. Wan, and F. L. Lewis, "Bayesian graphical games for synchronization in networks of dynamical systems," *IEEE Trans. Control Netw. Syst.*, vol. 7, no. 2, pp. 1028–1039, Jun. 2020.

[12] M. Liu, Y. Wan, V. G. Lopez, F. L. Lewis, G. A. Hewer, and K. Estabridis, "Differential graphical game with distributed global Nash solution," *IEEE Trans. Control Netw. Syst.*, vol. 8, no. 3, pp. 1371–1382, Sep. 2021.

[13] G. Tsaousoglou, "Correlated equilibrium power flow in distribution networks using graphical game theory," *IEEE Trans. Smart Grid*, vol. 14, no. 4, pp. 2621–2629, Jul. 2023.

[14] L. Li, C. Langbort, and J. Shamma, "An LP approach for solving two-player zero-sum repeated Bayesian games," *IEEE Trans. Autom. Control*, vol. 64, no. 9, pp. 3716–3731, Sep. 2019.

[15] V. Mittal, S. Maghsudi, and E. Hossain, "Distributed cooperation under uncertainty in drone-based wireless networks: A Bayesian coalitional game," *IEEE Trans. Mobile Comput.*, vol. 22, no. 1, pp. 206–221, Jan. 2023.

[16] X. Chen, J. Gao, D. Ge, and Z. Wang, "Bayesian dynamic learning and pricing with strategic customers," *Prod. Oper. Manag.*, vol. 31, no. 8, pp. 3125–3142, 2022.

[17] C. Eksin and A. Ribeiro, "Distributed fictitious play in potential games of incomplete information," in *Proc. 54th IEEE Conf. Decision Control (CDC)*, 2015, pp. 5190–5196.

[18] F. M. Aziz, L. Li, J. S. Shamma, and G. L. Stuber, "Smart jammer and LTE network strategies in an infinite-horizon zero-sum repeated game with asymmetric and incomplete information," 2019, *arXiv:1904.11184*.

[19] Y. Chen, L. Zhang, T. Merry, S. Amatya, W. Zhang, and Y. Ren, "When shall i be empathetic? The utility of empathetic parameter estimation in multi-agent interactions," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, 2021, pp. 2761–2767.

[20] M. Tian, Z. Dong, and X. Wang, "Analysis of false data injection attacks in power systems: A dynamic Bayesian game-theoretic approach," *ISA Trans.*, vol. 115, pp. 108–123, Sep. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0019057821000136

[21] L. Huang and Q. Zhu, "A dynamic game framework for rational and persistent robot deception with an application to deceptive pursuit-evasion," *IEEE Trans. Autom. Sci. Eng.*, vol. 19, no. 4, pp. 2918–2932, Oct. 2022.

[22] Y. Xu, H. Yang, B. Jiang, and M. M. Polycarpou, "Multiplayer pursuit-evasion differential games with malicious pursuers," *IEEE Trans. Autom. Control*, vol. 67, no. 9, pp. 4939–4946, Sep. 2022.

[23] F. L. Lewis, D. Vrabie, and V. L. Syrmos, *Optimal Control*, 3rd ed. Hoboken, NJ, USA: Wiley, 2012.

[24] Y.-F. Luo, *Course of Game Theory*. Beijing, China: Tsinghua Univ. Press, 2007.

[25] M. Milinski, “Tit for tat in sticklebacks and the evolution of cooperation,” *Nature*, vol. 325, no. 6103, pp. 433–435, 1987.
[26] M. J. Osborne and A. Rubinstein, *A Course in Game Theory*. Cambridge, MA, USA: MIT Press, 1994.
[27] R. Adderson and Y.-J. Pan, “Terminal sliding mode control for the formation of a team of quadrotors and mobile robots,” in *Proc. IECON 47th Annu. Conf. IEEE Ind. Electron. Soc.*, 2021, pp. 1–6.

Yue Zhang received the B.S. degree in computational mathematics from the Huazhong University of Science and Technology, Wuhan, China, in 2021, where he is currently pursuing the Ph.D. degree in control theory and control engineering.

His research interests include multiagent systems, distributed control, and game theory.

Yan-Wu Wang (Senior Member, IEEE) received the B.S. degree in automatic control, the M.S. and Ph.D. degrees in control theory and control engineering from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 1997, 2000, and 2003, respectively.

She has been a Professor with the School of Artificial Intelligence and Automation, HUST, since 2009. She is currently with the Key Laboratory of Image Processing and Intelligent Control, Ministry of Education, China. Her research interests include hybrid systems, cooperative control, and multiagent systems with applications in the smart grid.

Dr. Wang was a recipient of several awards, including the First Prize of Hubei Province Natural Science in 2014, the First Prize of the Ministry of Education of China in 2005, and the Excellent Ph.D. Dissertation of Hubei Province, China, in 2004. In 2008, she was awarded the title of “New Century Excellent Talents” by the Chinese Ministry of Education.

Yan Lei received the B.S. degree in automatic control from Hohai University, Changzhou, China, in 2016, and the Ph.D. degree in control theory and control engineering from the Huazhong University of Science and Technology (HUST), Wuhan, China, in 2021.

From October 2019 to March 2020, he was a Visiting Scholar with the Research Center for Automatic Control of Nancy, Nancy, France. From June 2022 to May 2023, he was a Postdoctoral Research Fellow with the Department of Electrical Engineering, Yeungnam University, Gyeongsan, South Korea. Since March 2021, he has been a Postdoctoral Research Fellow with the School of Artificial Intelligence and Automation, HUST. His research interests include hybrid systems, multiagent systems, and event-triggered control.

Yun-Feng Luo was born in 1966. He received the Ph.D. degree in systems engineering from the Huazhong University of Science and Technology, Wuhan, China, in 1994.

He was a Postdoctoral Research Fellow with the Huazhong University of Science and Technology, in 1997, where he is currently a Professor with the School of Artificial Intelligence and Automation. His current research interests include decision analysis, game theory, and social choice theory.

Zhi-Wei Liu (Senior Member, IEEE) received the B.S. degree in information management and information system from Southwest Jiaotong University, Chengdu, China, in 2004, and the Ph.D. degree in control science and engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2011.

He is currently a Professor with the School of Artificial Intelligence and Automation, Huazhong University of Science and Technology. His current research interests include cooperative control and optimization of distributed network systems.