SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

October 2024, Vol. 67, Iss. 10, 202204:1–202204:22
https://doi.org/10.1007/s11432-023-4035-9

Fault-tolerant flocking control against multiple malicious agents under geometric configuration containment

Chencheng ZHANG[1,2], Hao YANG[1*] & Bin JIANG[1]

[1]*College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing* 211106, *China;*
[2]*Institute of Engineering and Technology (ENTEG), University of Groningen, Groningen* 9747 AG, *the Netherlands*

Received 9 August 2023/Revised 22 September 2023/Accepted 3 October 2023/Published online 25 September 2024

Abstract Flocking control in the presence of malicious decisions is a crucial research topic since agents in swarms are likely to suffer from security issues in practice. This paper considers a swarm containing malicious agents that deliberately alter their controller parameters and disrupt the swarm by causing collisions, divisions, and escapes. Two possible malicious scenarios are considered. The first is that one malicious agent acts selectively on a subset of its neighbors and the second is that multiple malicious agents are either scattered or clustered together. To achieve the flocking control objective: (1) containment conditions based on specific geometric configurations are established; (2) hierarchical fault-tolerant flocking control methods based on those containment conditions are proposed. Compared to existing research dealing with malicious agents in swarms, our approach aims to contain rather than remove them. This guarantees the completeness of the task and takes agents' motions into consideration. Simulations show the effectiveness of the proposed methods.

Keywords flocking control, fault-tolerant control, multiple malicious agents, geometric configuration

1 Introduction

Flocking is a typical self-organizing motion that extracts energy from the environment and generates directed motion. Flocking motion is widely observed in nature, i.e., fish display schooling [1] and bird gathering [2]. Inspired by these natural occurrences, many researchers have focused on the flocking control problem, applying their findings to multi-agent systems, mobile agents, or networks [3–5]. Most of them are based on the three heuristic rules: separation, alignment, and cohesion for the flocking model proposed by Reynolds in [6]. Recently, new results on flocking control for multi-agent systems constantly spring up. Ref. [7] presented an inference-based approach to the distributed and cooperative flocking control of aerial robot swarms. In [8], an overview of optimal flocking is presented.

In large-scale systems, complicated environments and redundant interconnections pose significant challenges to the security of a swarm or a networked system. The major objective of providing security in any networked system is to defend the system against a range of abnormal situations. Motivated by these challenges, many fault-tolerant control methods have been developed [9]. These include strategies for addressing actuator or sensor failures in flight systems, with numerous studies proposing solutions such as fault tolerant adaptive control and artificial neural network methods [10, 11]. Beyond physical layer faults, malicious behaviors at the decision-making layer of a system present even more complex challenges to deal with. Ref. [12] provided a survey on trust-based detection and isolation of malicious nodes in ad-hoc and sensor networks. In [13], some protocols are proposed for the identification and local containment of misbehaving or faulty nodes, followed by their eviction from vehicular networks. For a manned-unmanned vehicle swarm, some members may intentionally gain control of vehicles to sabotage the mission of the whole swarm [14].

To date, there has been limited research focusing on flocking control within smarms that include malicious agents. These agents threaten the primary objectives of flocking because they increase the risk

* Corresponding author (email: haoyang@nuaa.edu.cn)

of collisions and splits within the swarm. The aforementioned behaviors can manifest from the malicious decisions made by these agents. Motivated by this, in our previous work [15], we addressed the flocking control problem of a swarm against one subjective malicious agent. That malicious agent deliberately falsifies its controller parameters, disrupting the balance between attraction or repulsion forces among agents, potentially causing collisions, divisions, and agent dispersal within the swarm. To guarantee the integrity of the task and consider the motions of the agent, the malicious agent is supposed to be safely contained in a swarm. It should be noted that its abnormal behavior typically impacts all neighboring agents indiscriminately.

However, real-world scenarios often involve targeted attacks by malicious agents within a swarm. Usually, targets, attackers, and defenders simultaneously exist in combat engagement scenarios. For example, Ref. [16] considered scenarios when a Trojan is triggered in both targeted and untargeted settings. In [17], the crossfire attack is a moving target for the same N-server area. On the other hand, in large-scale swarms, the presence of multiple malicious agents can pose a significantly greater threat than a single agent, potentially leading to the destruction of the whole swarm. Some researchers have focused on dealing with malicious nodes or agents. Ref. [18] proposed an onion-peeling approach to defend against multiple malicious nodes. In [19], a hybrid strategy is developed to reach resilient consensus among cooperative agents in the directed networks, even when multiple Byzantine agents are present.

Motivated by the above analysis, this work aims to explore more realistic scenarios, including those involving malicious agents with targets and the presence of multiple malicious agents within a swarm. Under these scenarios, we aim to study how the malicious agents affect the whole swarm and how to achieve the flocking control goal without removing the malicious agents from the swarm. The main contributions of this work are as follows:

• In scenarios where a malicious agent selectively targets certain neighbors, a geometric flocking condition is established. This condition ensures that the forces acting on the malicious agent by its target and non-target neighbors are balanced and that there are specific desired distances between the malicious agent and its neighbors. Leveraging this condition, a hierarchical geometric configuration-based flocking control strategy is designed. Such a fault-tolerant flocking method, for the first time, enables the containment of a malicious agent with targets by its neighboring agents.

• For scenarios involving multiple malicious agents, whether scattered throughout the swarm or clustered together, the desired geometric configuration and hierarchical flocking control strategies are developed by comprehensively analyzing the forces acting on all malicious agents. Moreover, not only the containment of every single malicious agent but also the velocity consensus of all malicious agents is considered to avoid the split of the swarm.

The remainder of the paper is organized as follows: In Section 2, a model description and some previous work are given. Section 3 provides a containment analysis for the malicious agent with targets and the flocking control methods. Section 4 focuses on multiple malicious agents. The simulation results are presented in Section 5, followed by a conclusion in Section 6.

2 Preliminaries

Notations. Let $\mathbf{1}_n$ denote the $n \times 1$ column vector of all ones. Let $|a|_1$ denote the 1-norm and $|a|$ denote the Euclidean-norm of a, respectively. Let $\mathrm{sgn}(a)$ be the signum function of a. Let $\mathrm{diag}(a_1, \ldots, a_p)$ be the diagonal matrix with diagonal entries a_1 to a_p. Let $\lambda_{\min}(\cdot)$ denote the minimum eigenvalue of a square real matrix with real eigenvalues. Let $\otimes$ be the Kronecker matrix product.

2.1 Flocking of a swarm

Consider a swarm of N agents whose dynamics take the form

$$\begin{cases} \dot{x}_i = v_i, \\ \dot{v}_i = u_i, \quad i \in \mathcal{V}, \end{cases} \tag{1}$$

where $x_i \in \Re^m$, $v_i \in \Re^m$ and $u_i \in \Re^m$ denote the position, the velocity, and the control (acceleration) input of agent i for $i \in \mathcal{V}$ with $\mathcal{V} \triangleq \{1, \ldots, N\}$. Define $x_{ij} \triangleq x_i - x_j$ as the relative position between agents i and j for $i, j \in \mathcal{V}$.

The communication topology between agents in swarm (1) is modeled by an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ that consists of a set of vertices $\mathcal{V}$ and a set of edges $\mathcal{E} \triangleq \{(i,j)|i,j \in \mathcal{V}, i \neq j\}$. Vertex $i \in \mathcal{V}$ represents agent i, and edge $(i,j) \in \mathcal{E}$ implies that agents i and j can interact with each other and are unordered. An undirected path between vertices i and j is a sequence of unordered edges, $(i,k_1),(k_1,k_2),\ldots,(k_l,j)$ with distinct vertices k_p, $p = 1,2,\ldots,l$. If there exists an undirected path between vertices i and j, the two vertices are said to be connected; otherwise, they are unconnected. An undirected graph is called connected if any two distinct vertices in the graph are connected. The Laplacian matrix of graph $\mathcal{G}$ is denoted by L. Define R as the sensing radius of each agent, which indicates that two agents can interact only if the distance between them is smaller than R, i.e., if $0 < |x_{ij}| < R$, then $(i,j) \in \mathcal{E}$; otherwise, $(i,j) \notin \mathcal{E}$. Agent j is called a neighbor of agent i if $(i,j) \in \mathcal{E}$. Define $\mathcal{N}(i) \triangleq \{j \in \mathcal{V} : (i,j) \in \mathcal{E}, i \neq j\}$ as the set of neighbors of agent i in $\mathcal{G}$. Note that the following study can be applied to the case that the communication topology is considered static as well.

In the following, the definition of the flocking control objective and the classic method to achieve the objective are presented.

Definition 1 (Flocking control objective)**.** The flocking control objective of the swarm (1) is said to be achieved if all the agents tend to a common speed and approach a fixed configuration without collision, i.e., $\lim_{t\to\infty} \dot{x}_{ij} = \lim_{t\to\infty} v_i - v_j = 0$, $\forall i,j \in \mathcal{V}$; $0 < |x_{ik}(t)| < R$, $t \geqslant 0$, $\forall i \in \mathcal{V}$, $k \in \mathcal{N}(i)$.

To achieve the above flocking control objective, a conventional flocking control law is designed as [3]

$$u_i = -\sum_{j\in\mathcal{N}(i)} (v_i - v_j) - \sum_{j\in\mathcal{N}(i)} \nabla_{x_i} V_{ij}(|x_{ij}|), \quad i \in \mathcal{V}, \tag{2}$$

where the first term corresponds to the desired velocity alignment, and the second term is the gradient of a potential function V_{ij}. Note that many existing potential functions with different forms can be applied here in the normal case, for example, the bounded potential function proposed in [20]

$$V_{ij}(|x_{ij}|) \triangleq \underbrace{\frac{R^2 - |x_{ij}|^2}{|x_{ij}|^2 + \frac{R^2}{E}}}_{V_{rij}} + \underbrace{\frac{|x_{ij}|^2}{R^2 - |x_{ij}|^2 + \frac{R^2}{E}}}_{V_{aij}}, \quad 0 \leqslant |x_{ij}| \leqslant R, \tag{3}$$

where E is a positive constant. V_{ij} satisfies the following properties:

- $V_{ij}(|x_{ij}|) = E$ when $|x_{ij}| = 0$ or $|x_{ij}| = R$;
- $\frac{\partial V_{ij}(|x_{ij}|)}{\partial(|x_{ij}|)} < 0$ when $|x_{ij}| \in (0,\delta)$ and $\frac{\partial V_{ij}(|x_{ij}|)}{\partial(|x_{ij}|)} > 0$ when $|x_{ij}| \in (\delta, R)$, where $\delta \triangleq \frac{\sqrt{2}R}{2}$.

Physically, the potential can be divided into $V_{ij} \triangleq V_{aij} + V_{rij}$ where V_{aij} and V_{rij} can be viewed as potentials of attraction and repulsion of agent i with respect to agent j, respectively. Obviously, V_{ij} reaches its minimum when $|x_{ij}| = \delta$. In the unique distance δ, it holds that $\nabla_{x_i} V_{aij}(\delta) + \nabla_{x_i} V_{rij}(\delta) = 0$. In normal case, one can choose $E > \bar{Q} \triangleq \frac{1}{2}\sum_{i\in\mathcal{V}} v_i^{\mathrm{T}}(0)v_i(0) + \frac{N(N-1)}{2}\max_{i,j\in\mathcal{V}}\{\bar{V}_{ij}(|x_{ij}(0)|)\}$ where $\bar{V}_{ij}(|x_{ij}|) \triangleq \frac{R^2-|x_{ij}|^2}{|x_{ij}|^2} + \frac{|x_{ij}|^2}{R^2-|x_{ij}|^2}$. This makes the potential between any two agents sufficiently large when the distance between them is equal to 0 or R, and thus avoids the collision while preserving the connectivity [20]. In the sequel, E will be chosen sufficiently large (i.e., larger than the initial energy functions built in the following sections) to avoid the collision and preserve the connectivity when applying the potential function V_{ij} in the control design.

2.2 A hierarchical flocking control against one malicious agent

In this subsection, we recall a hierarchical geometric configuration-based flocking control method for a swarm with only one malicious agent, which has been proposed in our previous work [15]. In the control architecture, the malicious agent is in Layer 1, all its neighbors are in Layer 2, and other agents in the swarm are in Layer 3. Define $\mathcal{V}_l$ as the set of agents in Layer 2, $\mathcal{V}_f$ as the set of agents in Layer 3, and $\mathcal{V}_g \triangleq \{i_f\} + \mathcal{N}(i_f)$ as the set of agents in Layers 1 and 2. The control architecture is shown in Figure 1.

In Layer 1, the malicious agent is denoted as $i_f \in \mathcal{V}$. It intentionally falsifies controller parameters such that

$$u_{i_f} = -k_v \sum_{j\in\mathcal{N}(i_f)} (v_{i_f} - v_j) - \sum_{j\in\mathcal{N}(i_f)} \nabla_{x_{i_f}} \tilde{V}_{i_f j}(|x_{i_f j}|), \tag{4}$$

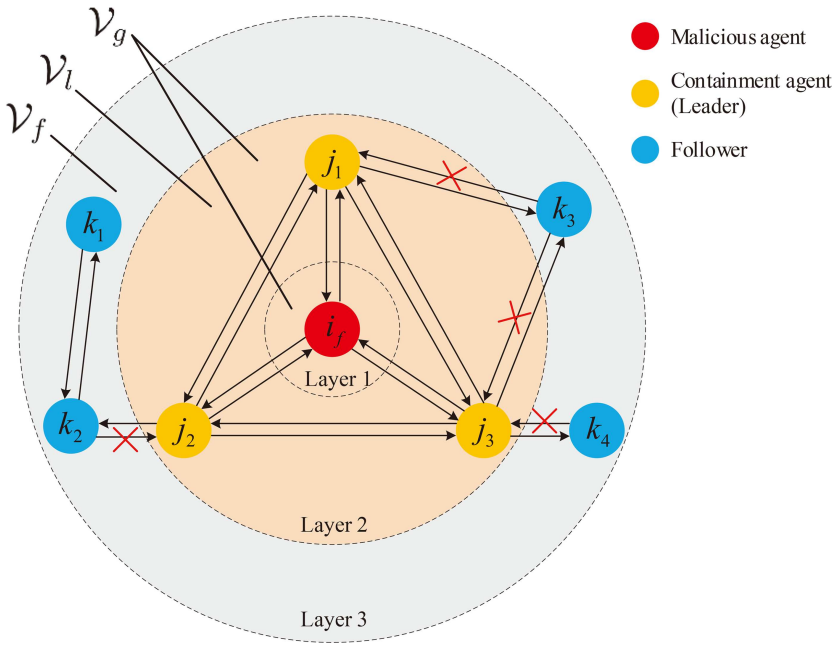

Figure 1 (Color online) Illustration of the hierarchical fault-tolerant control architecture [15].

where

$$\tilde{V}_{i_f j} \triangleq k_a V_{a i_f j}(|x_{i_f j}|) + k_r V_{r i_f j}(|x_{i_f j}|), \tag{5}$$

with $k_v < 1$, k_a, and k_r being three unknown bounded parameters. They represent the efficacy of the velocity consensus, and the strength of the attractive and repulsive force, respectively. $|k_v| \leqslant \bar{k}_v$, $|k_a| \leqslant \bar{k}_a$, $|k_r| \leqslant \bar{k}_r$ for $\bar{k}_v, \bar{k}_a, \bar{k}_r > 0$. Compared to the normal controller in (2), the attraction/repulsion effort acting on agent i_f from each of its neighbors is out of balance under the distance δ. Under this controller, especially when k_a or k_r is extremely large, the malicious agent i_f will run away from the agents around it to split the swarm or collide with the agents around it to ruin the swarm.

Definition 2. The malicious agent i_f is said to be contained if $\dot{v}_{i_f} = u_{i_f} = 0$ and $|x_{i_f j}| = \bar{\delta}_{i_f j}$ where $0 < \bar{\delta}_{i_f j} < R$ is a designable expected distance between agent i_f and its neighbor $j \in \mathcal{N}(i_f)$.

The malicious agent can be contained under the conditions shown in Lemma 1.

Lemma 1 ([15])**.** Consider the swarm (1) with malicious agent $i_f \in \mathcal{V}$ under controller (4) and (5). Suppose that $v_{i_f} - v_j = 0$, $\forall j \in \mathcal{N}(i_f)$. If

$$|x_{i_f j}| = \bar{\delta}, \ \forall j \in \mathcal{N}(i_f), \tag{6}$$

$$\sum_{j \in \mathcal{N}(i_f)} x_{i_f j} = 0, \tag{7}$$

where $0 < \bar{\delta} < R$, then $\dot{v}_{i_f} = u_{i_f} = 0$.

The assumptions on this method are provided here. For more details, please see [15].

Assumption 1. The unknown parameters k_v, k_a, k_r are bounded by $|k_v| \leqslant \bar{k}_v$, $|k_a| \leqslant \bar{k}_a$, $|k_r| \leqslant \bar{k}_r$.

Assumption 2. The initial undirected graph $\mathcal{G}'$ is connected, where $\mathcal{G}' \triangleq (\mathcal{V}', \mathcal{E}')$ consists of the set of vertices $\mathcal{V}' \triangleq \mathcal{V} - \{i_f\}$ and the set of edges $\mathcal{E}' \triangleq \{(i,j) | i,j \in \mathcal{V}', |x_{ij}| < R, i \neq j\}$.

Assumption 3. At the initial time, there are at least two neighbors of the malicious agent.

Assumption 4. At the initial time, any two agents in $\mathcal{N}(i_f)$ are neighbors.

Please note that Assumption 1 is not required if the unbounded potential functions are applied in this research. For example, the bounded potential function (3) can be replaced by the unbounded one $V_{ij}(|x_{ij}|) \triangleq \frac{1}{|x_{ij}|^2} + \frac{1}{(R-|x_{ij}|)^2}, 0 \leqslant |x_{ij}| \leqslant R$ meaning the design of V_{ij} does not depend on the bounds of these parameters.

Based on Conditions (6) and (7) of Lemma 1, let $\sum_{j \in \mathcal{N}(i_f)} x^*_{i_f j} = 0$ and $|x^*_{j i_f}| = \bar{\delta} < R/2$ where $x^*_{i_f j}$ denotes the desired displacement between agents i_f and $j \in \mathcal{N}(i_f)$.

In Layer 2, the controller of agent $j \in \mathcal{N}(i_f)$ is designed as follows:

$$u_j = -\kappa_v \sum_{p\in\mathcal{N}(j)\cap\mathcal{V}_l}(v_j - v_p) - \kappa_x \sum_{p\in\mathcal{N}(j)\cap\mathcal{V}_l} \nabla_{x_j}\hat{V}_{jp}(x_{jp}) - C_{i_f}\hat{k}, \tag{8}$$

where $k_v, k_x \geqslant 1$. $C_{i_f} \triangleq (\sum_{j\in\mathcal{N}(i_f)}(v_{i_f} - v_j), \sum_{j\in\mathcal{N}(i_f)} \nabla_{x_{i_f}} V_{ai_fj}(|x_{i_fj}|), \sum_{j\in\mathcal{N}(i_f)} \nabla_{x_{i_f}} V_{ri_fj}(|x_{i_fj}|))$. $v_{i_f}^F$ and C_v^F are the filtered functions of v_{i_f} and C_v under low-pass first-order filters, respectively. The estimate $\hat{k} \triangleq (\hat{k}_v, \hat{k}_a, \hat{k}_r)^{\mathrm{T}}$ of the unknown parameter $k \triangleq (k_v, k_a, k_r)^{\mathrm{T}}$ satisfies the following adaptive update law:

$$\dot{\hat{k}} = -\Gamma_k C_{i_f}^{\mathrm{T}} \sum_{j\in\mathcal{N}(i_f)}(v_j - v_{i_f}) - \Gamma_k (C_{i_f}^F)^{\mathrm{T}}(C_{i_f}^F\hat{k} + v_{i_f} - v_{i_f}^F), \tag{9}$$

where Γ_k is the positive-definite gain matrix. The non-negative potential function $\hat{V}_{ij}(x_{ij})$ satisfies the following properties that

(1) $\hat{V}_{ij}$ attains its unique minimum and $\frac{\partial \hat{V}_{ij}}{\partial |x_{ij} - x_{ij}^*|} = 0$ when $x_{ij} = x_{ij}^*$;

(2) $\hat{V}_{ij} > \bar{H}$ when $|x_{ij}| = 0$ and $|x_{ij}| = R$ where $\bar{H}$ is a designable positive constant.

In Layer 3, a distributed adaptive controller is designed for agent $k \in \mathcal{V}_f$ as

$$\begin{aligned} u_k &= -\sum_{p\in\mathcal{N}(k)} \alpha_{kp}\mathrm{sgn}(v_k - v_p) - \sum_{p\in\mathcal{N}(k)} \nabla_{x_k} V_{kp}(|x_{kp}|), \\ \dot{\alpha}_{kp} &= \gamma_{kp}|v_k - v_p|_1,\ p \in \mathcal{N}(k), \end{aligned} \tag{10}$$

where α_{kp} is a varying gain with initial values $\alpha_{kp}(0) \geqslant 0$ and V_{kp} is defined in (2). γ_{kp} is a positive constant and $\gamma_{kp} = \gamma_{pk}$

Lemma 2 ([15])**.** Consider the swarm (1) satisfying Assumptions 1–4 with malicious agent $i_f \in \mathcal{V}$ under controller (4) and (5). The malicious agent is contained, and the flocking control objective is achieved by applying controller (8) along with parameter estimate update law (9) to agents in $\mathcal{V}_l$ and controller (10) to agents in $\mathcal{V}_f$.

3 Flocking control against a malicious agent with targets

In our former work [15], the malicious agent is considered to act indiscriminately on all its neighbors. However, in practice, the malicious agent may focus on certain target agents. Therefore, before considering the multiple malicious agents' case, we first explore the scenario where the malicious agent acts selectively on a subset of its neighbors. To handle such a case, an extended hierarchical geometric configuration control method is provided.

Consider a malicious agent $i_f \in \mathcal{V}$ with specific targets, whose controller is

$$u_{i_f} = -k_v \sum_{j\in a(i_f)}(v_{i_f} - v_j) - \sum_{j\in a(i_f)} \nabla_{x_{i_f}} \tilde{V}_{i_fj}(|x_{i_fj}|) - \sum_{j\in\mathcal{N}(i_f)-a(i_f)} \nabla_{x_{i_f}} V_{i_fj}(|x_{i_fj}|), \tag{11}$$

where $a(i_f) \subset \mathcal{N}(i_f)$ is the set of agent i_f's targets. V_{i_fj} is defined in (3), k_v and $\tilde{V}_{i_fj}$ are defined in (5). Define $C_1 k \triangleq k_v \sum_{j\in a(i_f)}(v_{i_f} - v_j) + \sum_{j\in a(i_f)} \nabla_{x_{i_f}} \tilde{V}_{i_fj}(|x_{i_fj}|)$ where $C_1 \triangleq (\sum_{j\in a(i_f)}(v_{i_f} - v_j), \sum_{j\in a(i_f)} \nabla_{x_{i_f}} V_{ai_fj}, \sum_{j\in a(i_f)} \nabla_{x_{i_f}} V_{ri_fj})$ and $k \triangleq (k_v, k_a, k_r)^{\mathrm{T}}$. Define $f_1 \triangleq \sum_{j\in\mathcal{N}(i_f)-a(i_f)} \nabla_{x_{i_f}} V_{i_fj}$.

The flocking problem against this malicious agent is formulated as follows.

Problem $\mathcal{F}_1$. Consider the swarm (1) satisfying Assumptions 2 and 3 with a malicious agent $i_f \in \mathcal{V}$ under controller (11). Design u_i, $i \in \mathcal{V}'$ such that

(1) $\lim_{t\to\infty}(v_i - v_j) = 0,\ \forall i, j \in \mathcal{V}$, i.e., all the agents converge to the same velocity;

(2) The swarm (1) asymptotically converges to a fixed geometric configuration, under which

• $u_{i_f} = 0$ and $|x_{i_fj}| = \bar{\delta}_{i_fj}$, $\forall j \in \mathcal{N}(i_f)$ where $0 < \bar{\delta}_{i_fj} < R$, i.e., the malicious agent i_f is contained;

• $|x_{ij}| = \tilde{\delta}_{ij}$, $\forall i \in \mathcal{V}', j \in \mathcal{N}(i)$ where $0 < \tilde{\delta}_{ij} < R$, i.e., the normal agents are connected with their neighbors;

(3) $|x_{ij}(t)| \neq 0$ for $t \geqslant 0$, $\forall i, j \in \mathcal{V}$ and $i \neq j$, i.e., no collision occurs.

Before designing the controllers to achieve the flocking goal, a desired configuration to contain the malicious agent with targets is first investigated.

Lemma 3. Consider the swarm (1) with malicious agent $i_f \in \mathcal{V}$ under controller (11). Suppose that $v_{i_f} - v_j = 0$, $\forall j \in a(i_f)$. If

$$|x_{i_f j}| = \delta,\ \forall j \in \mathcal{N}(i_f) - a(i_f), \tag{12}$$

$$|x_{i_f j}| = \bar{\delta},\ \forall j \in a(i_f), \tag{13}$$

$$\sum_{j \in a(i_f)} x_{i_f j} = 0, \tag{14}$$

where $\delta, \bar{\delta} \in (0, R)$, then $\dot{v}_{i_f} = u_{i_f} = 0$.

Proof. Since $V_{i_f j}$ in (3) is symmetric with respect to $x_{i_f j}$ for $j \in \mathcal{N}(i_f)$, it holds that

$$\sum_{j \in \mathcal{N}(i_f) - a(i_f)} \nabla_{x_{i_f}} V_{i_f j}(|x_{i_f j}|) = \sum_{j \in \mathcal{N}(i_f) - a(i_f)} \nabla_{x_{i_f j}} V_{i_f j}(|x_{i_f j}|) = \sum_{j \in \mathcal{N}(i_f) - a(i_f)} \frac{\partial V_{i_f j}(|x_{i_f j}|)}{\partial |x_{i_f j}|} \cdot \frac{x_{i_f j}}{|x_{i_f j}|}.$$

According to the properties of $V_{i_f j}$, $\partial V_{i_f j}(|x_{i_f j}|)/\partial |x_{i_f j}| = 0$ when $|x_{i_f j}| = \delta$, $j \in \mathcal{N}(i_f) - a(i_f)$. Therefore, under condition (12), it holds that

$$\sum_{j \in \mathcal{N}(i_f) - a(i_f)} \nabla_{x_{i_f}} V_{i_f j}(|x_{i_f j}|) = 0. \tag{15}$$

Since $\tilde{V}_{i_f j}$ in (5) is symmetric with respect to $x_{i_f j}$ for $j \in a(i_f)$, it holds that

$$\sum_{j \in a(i_f)} \nabla_{x_{i_f}} \tilde{V}_{i_f j}(|x_{i_f j}|) = \sum_{j \in a(i_f)} \nabla_{x_{i_f j}} \tilde{V}_{i_f j}(|x_{i_f j}|) = \sum_{j \in a(i_f)} \frac{\partial \tilde{V}_{i_f j}(|x_{i_f j}|)}{\partial |x_{i_f j}|} \cdot \frac{x_{i_f j}}{|x_{i_f j}|}.$$

According to conditions (13) and (14), it yields that

$$\sum_{j \in a(i_f)} \nabla_{x_{i_f}} \tilde{V}_{i_f j}(|x_{i_f j}|) = s_a \cdot \left.\frac{\partial \tilde{V}_{i_f j}(|x_{i_f j}|)}{\partial |x_{i_f j}|}\right|_{|x_{i_f j}| = \bar{\delta}} \cdot \frac{\sum_{j \in a(i_f)} x_{i_f j}}{\bar{\delta}} = 0, \tag{16}$$

where s_a is the number of agents in set $a(i_f)$. Suppose that $v_{i_f} - v_j = 0$, $\forall j \in a(i_f)$. Under controller (11) and conditions (12)–(14), it holds that

$$\dot{v}_{i_f} = u_{i_f} = -k_v \sum_{j \in a(i_f)} (v_{i_f} - v_j) - \sum_{j \in \mathcal{N}(i_f) - a(i_f)} \nabla_{x_{i_f}} V_{i_f j}(|x_{i_f j}|) - \sum_{j \in a(i_f)} \nabla_{x_{i_f}} \tilde{V}_{i_f j}(|x_{i_f j}|) = 0.$$

This completes the proof.

Remark 1. Conditions (12)–(14) in Lemma 3 provide a desired geometric configuration that can contain the malicious agent with targets. The attraction and repulsion forces exerted on the malicious agent from its non-target neighbors are balanced once the distance condition (12) is satisfied. Under conditions (13) and (14), the configuration of the malicious agent and its targets is a regular polygon with the malicious agent being the center and its targets being vertexes. This balances physically the forces acting on the malicious agent from all its targets. Figure 2 illustrates a desired configuration where the malicious agent has four neighbors and agents $j_1 - j_3$ are targets.

To deal with the unknown parameters in the controller of malicious agents with targets, let

$$\dot{v}_{i_f}^F = -a v_{i_f}^F + a v_{i_f},\quad v_{i_f}^F(0) = v_{i_f}(0), \tag{17}$$

$$\dot{f}_1^F = -a f_1^F + f_1,\quad f_1^F(0) = 0, \tag{18}$$

$$\dot{C}_1^F = -a C_1^F + C_1,\quad C_1^F(0) = 0, \tag{19}$$

where $v_{i_f}^F$, f_1^F and C_1^F are the filtered functions of v_{i_f}, f_1 and C_1, respectively. Based on these functions, we design the adaptive update law of parameter k's estimate as

$$\dot{\hat{k}} = -\Gamma_k C_1^{\mathrm{T}} \sum_{j \in \mathcal{N}(i_f)} (v_j - v_{i_f}) - \Gamma_k (C_1^F)^{\mathrm{T}} (C_1^F \hat{k} - v_{i_f}^F + v_{i_f} + f_1^F). \tag{20}$$

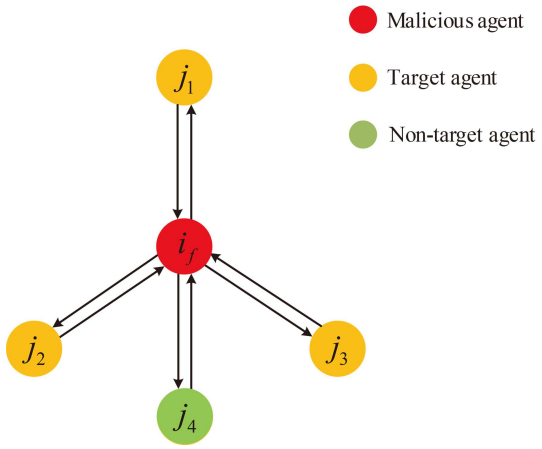

Figure 2 (Color online) Illustration of the configuration to contain the malicious agent with targets.

According to Lemma 3, let $\sum_{j\in a(i_f)} x^*_{ji_f} = 0$ and $|x^*_{ji_f}| = \bar{\delta} < R/2$, $\forall j \in a(i_f)$. Let $|x^*_{ji_f}| = \delta$ $\forall j \in \mathcal{N}(i_f) - a(i_f)$. Design the geometric configuration controller u_j, $j \in \mathcal{N}(i_f)$,

$$\begin{aligned} u_j =& -\kappa_v \sum_{p\in\mathcal{N}(j)\cap\mathcal{V}_g} (v_j - v_p) - \kappa_x \sum_{p\in\mathcal{N}(j)\cap\mathcal{V}_g} \nabla_{x_j}\hat{V}_{jp}(x_{jp}) \\ & - \sum_{b\in\mathcal{N}(i_f)-a(i_f)} \nabla_{x_{i_f}} V_{i_f b}(|x_{i_f b}|) - C_1\hat{k}, \end{aligned} \tag{21}$$

where κ_v, κ_x and $\hat{V}_{ij}$ are defined in (8). $V_{i_f b}$ is defined in (3) and C_1 is defined in (11).

Theorem 1. Consider the swarm (1) satisfying Assumptions 1–4 with malicious agent $i_f \in \mathcal{V}$ under controller (11). Supposing $s_a \geqslant 2$, Problem $\mathcal{F}_1$ is solved by applying controller (21) along with parameter estimate update law (20) to agents in $\mathcal{V}_l$ and controller (10) to agents in $\mathcal{V}_f$.

Proof. Denote $x_f \triangleq (x_{i_f}^{\mathrm{T}}, x_{j_1}^{\mathrm{T}}, \ldots, x_{j_s}^{\mathrm{T}})^{\mathrm{T}}$, $v_f \triangleq (v_{i_f}^{\mathrm{T}}, v_{j_1}^{\mathrm{T}}, \ldots, v_{j_s}^{\mathrm{T}})^{\mathrm{T}}$ for $j_k \in \mathcal{N}(i_f)$, $k \in \{1, \ldots, s\}$. Define an energy function $H(x_f, v_f)$ as

$$\begin{aligned} H(x_f, v_f) \triangleq& \kappa_x \sum_{j\in\mathcal{N}(i_f)} \hat{V}_{ji_f} + \frac{\kappa_x}{2} \sum_{j\in\mathcal{N}(i_f)} \sum_{i\in\mathcal{N}(i_f)\cap\mathcal{N}(j)} \hat{V}_{ji} + \frac{1}{2} \sum_{j\in\mathcal{N}(i_f)} (v_j - v_{i_f})^{\mathrm{T}}(v_j - v_{i_f}) \\ & + \frac{1}{2}\tilde{k}^{\mathrm{T}}\Gamma_k^{-1}\tilde{k}, \end{aligned} \tag{22}$$

where $\tilde{k} \triangleq (\tilde{k}_v, \tilde{k}_a, \tilde{k}_r)^{\mathrm{T}}$ with $\tilde{k}_v \triangleq k_v - \hat{k}_v, \tilde{k}_a \triangleq k_a - \hat{k}_a, \tilde{k}_r \triangleq k_r - \hat{k}_r$. Note that $\dot{\hat{V}}_{ji} = \dot{x}_{ji}\nabla_{x_{ji}}\hat{V}_{ji}$. The time derivative of $H(x_f, v_f)$ is

$$\begin{aligned} \dot{H}(x_f, v_f) =& \kappa_x \sum_{j\in\mathcal{N}(i_f)} (v_j^{\mathrm{T}} - v_{i_f}^{\mathrm{T}})\nabla_{x_j}\hat{V}_{ji_f} + \frac{\kappa_x}{2} \sum_{j\in\mathcal{N}(i_f)} \sum_{i\in\mathcal{N}(i_f)\cap\mathcal{N}(j)} (v_j^{\mathrm{T}} - v_i^{\mathrm{T}})\nabla_{x_j}\hat{V}_{ji} \\ & + \sum_{j\in\mathcal{N}(i_f)} (v_j - v_{i_f})^{\mathrm{T}}(\dot{v}_j - \dot{v}_{i_f}) - \tilde{k}^{\mathrm{T}}\Gamma_k^{-1}\dot{\hat{k}}. \end{aligned}$$

Considering the estimator (20), it holds that

$$-v_{i_f}^F + v_{i_f} = -C_{i_f}^F k - f_1^F. \tag{23}$$

Applying equation (23) and controller (21), one has

$$\begin{aligned} \dot{H}(x_f, v_f) =& \kappa_x \sum_{j\in\mathcal{N}(i_f)} (v_j^{\mathrm{T}} - v_{i_f}^{\mathrm{T}})\nabla_{x_j}\hat{V}_{ji_f} + \frac{\kappa_x}{2} \sum_{j\in\mathcal{N}(i_f)} \sum_{i\in\mathcal{N}(i_f)\cap\mathcal{N}(j)} (v_j^{\mathrm{T}} - v_i^{\mathrm{T}})\nabla_{x_j}\hat{V}_{ji} \\ & + \sum_{j\in\mathcal{N}(i_f)} (v_j - v_{i_f})^{\mathrm{T}} \Bigg(- \kappa_v \sum_{p\in\mathcal{N}(j)\cap\mathcal{V}_g} (v_j - v_p) - \kappa_x \sum_{p\in\mathcal{N}(j)\cap\mathcal{V}_l} \nabla_{x_j}\hat{V}_{jp}(x_{jp}) \end{aligned}$$

$$
\begin{aligned}
&- C_{i_f}(\hat{k}-k)\Big) + \tilde{k}^{\mathrm{T}} C_{i_f}^{\mathrm{T}} \sum_{j\in\mathcal{N}(i_f)} (v_j - v_{i_f}) + \tilde{k}^{\mathrm{T}} (C_{i_f}^F)^{\mathrm{T}} (C_{i_f}^F \hat{k} - v_{i_f} + v_{i_f}^F + f_1^F) \\
=&\kappa_x \sum_{j\in\mathcal{N}(i_f)} (v_j^{\mathrm{T}} - v_{i_f}^{\mathrm{T}}) \nabla_{x_j} \hat{V}_{ji_f} + \frac{\kappa_x}{2} \sum_{j\in\mathcal{N}(i_f)} \sum_{i\in\mathcal{N}(i_f)\cap\mathcal{N}(j)} (v_j^{\mathrm{T}} - v_i^{\mathrm{T}}) \nabla x_j \hat{V}_{ji} \\
&+ \sum_{j\in\mathcal{N}(i_f)} (v_j - v_{i_f})^{\mathrm{T}} \Bigg(- \kappa_v (v_j - v_{i_f}) - \kappa_v \sum_{p\in\mathcal{N}(i_f)\cap\mathcal{N}(j)} (v_j - v_p) \\
&- \kappa_x \nabla_{x_j} \hat{V}_{ji_f} - \kappa_x \sum_{i\in\mathcal{N}(i_f)\cap\mathcal{N}(j)} \nabla_{x_j} \hat{V}_{ji} \Bigg) - \tilde{k}^{\mathrm{T}} (C_{i_f}^F)^{\mathrm{T}} C_{i_f}^F \tilde{k}. \qquad (24)
\end{aligned}
$$

It follows from the fact that $x_{ji} = -x_{ij}$ and $x_{ji} - x_{ji}^* = -(x_{ij} - x_{ij}^*)$ that

$$
\frac{\partial \hat{V}_{ji}}{\partial x_{ji}} = \frac{\partial \hat{V}_{ji}}{\partial x_j} = -\frac{\partial \hat{V}_{ij}}{\partial x_i}.
$$

Therefore, it holds that

$$
\begin{aligned}
&\frac{\kappa_x}{2} \sum_{j\in\mathcal{N}(i_f)} \sum_{i\in\mathcal{N}(i_f)\cap\mathcal{N}(j)} (v_j^{\mathrm{T}} - v_i^{\mathrm{T}}) \nabla_{x_j} \hat{V}_{ji} \\
&= \frac{\kappa_x}{2} \sum_{j\in\mathcal{N}(i_f)} \sum_{i\in\mathcal{N}(i_f)\cap\mathcal{N}(j)} \big((v_j^{\mathrm{T}} - v_{i_f}^{\mathrm{T}}) \nabla_{x_j} \hat{V}_{ji} + (v_i^{\mathrm{T}} - v_{i_f}^{\mathrm{T}}) \nabla_{x_i} \hat{V}_{ij} \big) \\
&= \kappa_x \sum_{j\in\mathcal{N}(i_f)} (v_j^{\mathrm{T}} - v_{i_f}^{\mathrm{T}}) \sum_{i\in\mathcal{N}(i_f)\cap\mathcal{N}(j)} \nabla_{x_j} \hat{V}_{ji}. \qquad (25)
\end{aligned}
$$

Combining (24) and (25) yields that

$$
\begin{aligned}
\dot{H}(x_f, v_f) =& - \kappa_v \sum_{j\in\mathcal{N}(i_f)} (v_j - v_{i_f})^{\mathrm{T}} (v_j - v_{i_f}) - \tilde{k}^{\mathrm{T}} (C_{i_f}^F)^{\mathrm{T}} C_{i_f}^F \tilde{k} \\
&- \frac{\kappa_v}{2} \sum_{j\in\mathcal{N}(i_f)} \sum_{p\in\mathcal{N}(i_f)\cap\mathcal{N}(j)} ((v_j - v_{i_f}) - (v_p - v_{i_f}))^{\mathrm{T}} (v_j - v_p) \\
=& - \kappa_v \sum_{j\in\mathcal{N}(i_f)} (v_j - v_{i_f})^{\mathrm{T}} (v_j - v_{i_f}) - \frac{\kappa_v}{2} \sum_{j\in\mathcal{N}(i_f)} \sum_{p\in\mathcal{N}(i_f)\cap\mathcal{N}(j)} (v_j - v_p)^{\mathrm{T}} (v_j - v_p) \\
&- \tilde{k}^{\mathrm{T}} (C_{i_f}^F)^{\mathrm{T}} C_{i_f}^F \tilde{k}. \qquad (26)
\end{aligned}
$$

Therefore, $\dot{H}(x_f, v_f)$ is always nonpositive and $H(t) \leqslant H(0)$ for $t \geqslant 0$. From the definition of $H(t)$ in (22), it holds that $H(t) > \hat{V}_{ji}(t)$, $\forall i, j \in \mathcal{V}_g$. Thus, $\hat{V}_{ji}(t) < H(0)$ for $t \geqslant 0$. According to the property (2) of $\hat{V}$, $\hat{V}_{ji}(t) > \bar{H}$, $\forall j \in \mathcal{N}(i)$ when $|x_{ji}| = 0$ and $|x_{ji}| = R$. The designable constant $\bar{H}$ is chosen such that $\bar{H} > H(0)$. Thus, $\hat{V}_{ji}(t) > \bar{H} > H(0)$ when $|x_{ji}| = 0$ and $|x_{ji}| = R$. This is contradicted to $\hat{V}_{ji}(t) < H(0), \forall t \geqslant 0$. Hence, $|x_{ji}| \neq 0$ and $|x_{ji}(t)| \neq R, \forall t \geqslant 0$. This ensures collision is prevented, and no edge is lost between any two agents in $\mathcal{V}_g$.

Define the level set $\Omega_f \triangleq \{(x_f^{\mathrm{T}}, v_f^{\mathrm{T}})^{\mathrm{T}} \in R^{2(s+1)\times m} : H(t) \leqslant H(0), H(0) > 0\}$. By applying LaSalle's invariance principle, $(x_f^{\mathrm{T}}, v_f^{\mathrm{T}})^{\mathrm{T}}$ starting in Ω_f asymptotically converges to the largest invariant set inside the region $\mathcal{C} \triangleq \{(x_f^{\mathrm{T}}, v_f^{\mathrm{T}})^{\mathrm{T}} \in \Omega_f : \dot{H}(t) = 0\}$. According to (26), $\dot{H}(t) = 0$ holds if and only if $v_1 = v_2 = \cdots = v_{i_f}$ and $C_{i_f}^F \tilde{k} = 0$. This implies that the malicious agent and its neighbors converge toward the same velocity, i.e., $\lim_{t\to\infty}(v_i - v_j) = 0$, $\forall i, j \in \mathcal{V}_g$. And $\lim_{t\to\infty} C_{i_f}^F \tilde{k} = 0$. Moreover, according to (18), $\lim_{t\to\infty} C_{i_f}^F = C_{i_f}/a$ where $a > 0$. Thus $\lim_{t\to\infty} C_{i_f} \tilde{k} = 0$.

In the following, we consider the error system $\dot{v}_j - \dot{v}_{i_f} = u_j - u_{i_f}$ for $j \in \mathcal{N}(i_f)$ at the point $v_1 = v_2 = \cdots = v_{i_f}$. Obviously, it holds that $\dot{v}_j - \dot{v}_i = 0, \forall i, j \in \mathcal{V}_g$. Combining malicious controller (4) and controller (8), one has $\dot{v}_j - \dot{v}_{i_f} = -\kappa_v \sum_{j\in\mathcal{N}(i_f)} (v_{i_f} - v_j) - \sum_{p\in\mathcal{N}(j)\cap\mathcal{V}_g} (v_j - v_p) - \sum_{p\in\mathcal{N}(j)\cap\mathcal{V}_g} \nabla_{x_j} \hat{V}_{jp} + C_{i_f} \tilde{k}$. Note that $C_{i_f} \tilde{k} = 0$ at the point $v_1 = v_2 = \cdots = v_{i_f}$ $\forall j \in \mathcal{V}_g$. Thus, $\sum_{p\in\mathcal{N}(j)\cap\mathcal{V}_g} \nabla_{x_j} \hat{V}_{jp} = 0$. Define

$\psi_{\mathcal{G}_f}(x_f) \triangleq (\ldots, |x_{ij} - x_{ij}^*|, \ldots)^{\mathrm{T}}$ with $i, j \in \mathcal{V}_g$. Considering the error system in a compact form, one obtains that $R_{\mathcal{G}_f}^{\mathrm{T}}(x_f)\xi(x_f) = 0$ where $\xi(x_f) \triangleq (\ldots, \partial\hat{V}_{ij}/\partial|x_{ij} - x_{ij}^*|, \ldots)$ and $R_{\mathcal{G}_f}^{\mathrm{T}}(x_f) \triangleq \nabla x_f \psi_{\mathcal{G}_f}(x_f)$ is the rigidity matrix. Since $\mathrm{Rank}(R_{\mathcal{G}_f}(x_f)) = md - m(m+1)/2$ where m is the dimension and d is the vertex number of $\mathcal{V}_g$, it follows from the rigidity theory in [21] that $R_{\mathcal{G}_f}^{\mathrm{T}}(x_f)\xi(x_f) = 0$ is equivalent to $\xi(x_f) = 0$. From the property (1) of $\hat{V}_{ij}$, we can deduce that $\partial\hat{V}_{ij}/\partial|x_{ij} - x_{ij}^*| = 0$ is equivalent to $|x_{ij} - x_{ij}^*| = 0, \forall i, j \in \mathcal{V}_g$. Hence, it holds that $x_{ij} \to x_{ij}^*$ as $t \to \infty$, which means the desired configuration of $\mathcal{V}_g$ is approached. Also, it yields from controllers (4) and (8) that $\dot{v}_{i_f} = u_{i_f} \to 0$ and $\dot{v}_j = u_j \to 0$ for $j \in \mathcal{N}(i_f)$. According to the above analysis, the agents in $\mathcal{V}_g$ approach the desired configuration and converge toward the same velocity.

Let $\bar{\mathcal{G}}$ be the direct graph characterizing the information interactions among agents in $\mathcal{V}_f$ and the transmission from agents in $\mathcal{V}_l(\mathcal{N}(i_f))$ to agents in $\mathcal{V}_f$. If there is a directed path from agent $j \in \mathcal{N}(i_f)$ to agent $k \in \mathcal{V}_f$ in graph $\bar{\mathcal{G}}$, agent j is regarded a leader of agent k. Denote $\mathcal{L}(k)$ as the set of agent k's leaders. Define the energy function $\Upsilon(x, v)$ as

$$\begin{aligned}\Upsilon(x,v) \triangleq & H(x_f, v_f) + \sum_{i\in\mathcal{V}_f}\sum_{j\in\mathcal{L}(i)\cap\mathcal{N}(i)} V_{ij} + \frac{1}{2}\sum_{i\in\mathcal{V}_f}\sum_{p\in\mathcal{V}_f\cap\mathcal{N}(i)} s_{\mathcal{L}(i)}V_{ip} + \frac{1}{2}\sum_{i\in\mathcal{V}_f}\sum_{j\in\mathcal{L}(i)}(v_i - v_j)^{\mathrm{T}}(v_i - v_j) \\ & + \sum_{i\in\mathcal{V}_f}\sum_{j\in\mathcal{N}(i_f)\cap\mathcal{N}(i)} \frac{1}{2\gamma_{ij}}(\alpha_{ij} - \bar{\alpha})^2 + \sum_{i\in\mathcal{V}_f}\sum_{p\in\mathcal{V}_f\cap\mathcal{N}(i)} \frac{s_{\mathcal{L}(i)}}{4\gamma_{ip}}(\alpha_{ip} - \bar{\alpha})^2, \end{aligned} \tag{27}$$

where $s_{\mathcal{L}(i)}$ denotes the number of agents in set $\mathcal{L}(i)$. Constant $\bar{\alpha}$ will be designed later. Note that the graph of agents in Layer 3 is undirected, thus $s_{\mathcal{L}(i)} = s_{\mathcal{L}(j)}$, $\forall i \in \mathcal{V}_f, j \in \mathcal{N}(i) \cap \mathcal{V}_f$. Therefore, the derivative of Υ is

$$\begin{aligned}\dot{\Upsilon} = & \dot{H} + \sum_{i\in\mathcal{V}_f}\sum_{j\in\mathcal{L}(i)\cap\mathcal{N}(i)} \dot{x}_{ij}^{\mathrm{T}}\nabla_{x_{ij}}V_{ij} + \frac{1}{2}\sum_{i\in\mathcal{V}_f}\sum_{p\in\mathcal{V}_f\cap\mathcal{N}(i)} s_{\mathcal{L}(i)}\dot{x}_{ip}^{\mathrm{T}}\nabla_{x_{ip}}V_{ip} \\ & + \sum_{i\in\mathcal{V}_f}\sum_{j\in\mathcal{L}(i)}(v_i - v_j)^{\mathrm{T}}(\dot{v}_i - \dot{v}_j) + \sum_{i\in\mathcal{V}_f}\sum_{j\in\mathcal{N}(i_f)\cap\mathcal{N}(i)}(\alpha_{ij} - \bar{\alpha})|v_i - v_j|_1 \\ & + \sum_{i\in\mathcal{V}_f}\sum_{p\in\mathcal{V}_f\cap\mathcal{N}(i)} \frac{s_{\mathcal{L}(i)}}{2}(\alpha_{ip} - \bar{\alpha})|v_i - v_p|_1. \end{aligned}$$

Applying controller (10), one obtains that

$$\begin{aligned}\dot{\Upsilon} = & \dot{H} + \sum_{i\in\mathcal{V}_f}\sum_{j\in\mathcal{L}(i)\cap\mathcal{N}(i)} \dot{x}_{ij}^{\mathrm{T}}\nabla_{x_{ij}}V_{ij} + \frac{1}{2}\sum_{i\in\mathcal{V}_f}\sum_{j\in\mathcal{V}_f\cap\mathcal{N}(i)} s_{\mathcal{L}(i)}\dot{x}_{ij}^{\mathrm{T}}\nabla_{x_{ij}}V_{ij} \\ & + \sum_{i\in\mathcal{V}_f}\sum_{j\in\mathcal{L}(i)}(v_i - v_j)^{\mathrm{T}}\left(-\sum_{p\in\mathcal{N}(i)}\mathrm{sgn}(v_i - v_p) - \sum_{p\in\mathcal{N}(i)}\nabla_{x_{ip}}V_{ip} - u_j\right) \\ & + \sum_{i\in\mathcal{V}_f}\sum_{j\in\mathcal{N}(i_f)\cap\mathcal{N}(i)}(\alpha_{ij} - \bar{\alpha})|v_i - v_j|_1 + \sum_{i\in\mathcal{V}_f}\sum_{p\in\mathcal{V}_f\cap\mathcal{N}(i)} \frac{s_{\mathcal{L}(i)}}{2}(\alpha_{ip} - \bar{\alpha})|v_i - v_p|_1 \\ = & \dot{H}(x_f, v_f) + \sum_{i\in\mathcal{V}_f}\sum_{j\in\mathcal{L}(i)\cap\mathcal{N}(i)}(v_i - v_j)^{\mathrm{T}}\nabla_{x_{ij}}V_{ij} + \frac{1}{2}\sum_{i\in\mathcal{V}_f}\sum_{p\in\mathcal{V}_f\cap\mathcal{N}(i)} s_{\mathcal{L}(i)}\dot{x}_{ip}^{\mathrm{T}}\nabla_{x_{ip}}V_{ip} \\ & + \sum_{i\in\mathcal{V}_f}\sum_{j\in\mathcal{L}(i)}(v_i - v_j)^{\mathrm{T}}\Bigg(-\sum_{j\in\mathcal{N}(i_f)\cap\mathcal{N}(i)}\alpha_{ij}\mathrm{sgn}(v_i - v_j) - \sum_{p\in\mathcal{V}_f\cap\mathcal{N}(i)}\alpha_{ip}\mathrm{sgn}(v_i - v_p) \\ & - \sum_{j\in\mathcal{L}(i)\cap\mathcal{N}(i)}\nabla_{x_{ij}}V_{ij} - \sum_{p\in\mathcal{V}_f\cap\mathcal{N}(i)}\nabla_{x_{ip}}V_{ip} - u_j\Bigg) + \sum_{i\in\mathcal{V}_f}\sum_{j\in\mathcal{N}(i_f)\cap\mathcal{N}(i)}(\alpha_{ij} - \bar{\alpha})|v_i - v_j|_1 \\ & + \sum_{i\in\mathcal{V}_f}\sum_{p\in\mathcal{V}_f\cap\mathcal{N}(i)} \frac{s_{\mathcal{L}(i)}}{2}(\alpha_{ip} - \bar{\alpha})|v_i - v_p|_1. \end{aligned}$$

For convenience, label the agents in $\mathcal{N}(i_f)$ who have neighbors in $\mathcal{V}_f$ as 1 to ω. If there is a directed path from agent j, $j \in \{1, \ldots, \omega\}$ to some agents in $\mathcal{V}_f$, denote the set of these agents as $F(j)$. Note that

$F(j) \subseteq \mathcal{V}_f$, and the leaders of k, p are same if $k, p \in \mathcal{V}_f$ are neighbors. Therefore, it follows:

$$\begin{aligned}\dot{\Upsilon} =&\dot{H} + \frac{1}{2}\sum_{i\in\mathcal{V}_f}\sum_{p\in\mathcal{V}_f\cap\mathcal{N}(i)} s_{\mathcal{L}(i)}\dot{x}_{ip}^{\mathrm{T}}\nabla_{x_{ip}}V_{ip} - \sum_{i\in\mathcal{V}_f}\sum_{j\in\mathcal{N}(i_f)\cap\mathcal{N}(i)}\alpha_{ij}|v_i - v_j|_1\\ &-\sum_{j=1}^{\omega}\sum_{i\in F(j)}(v_i - v_j)^{\mathrm{T}}\sum_{p\in F(j)\cap\mathcal{N}(i)}\alpha_{ip}\mathrm{sgn}((v_i - v_j) - (v_p - v_j))\\ &-\sum_{j=1}^{\omega}\sum_{i\in F(j)}(v_i - v_j)^{\mathrm{T}}\sum_{p\in\mathcal{V}_f\cap\mathcal{N}(i)}\nabla_{x_{ip}}V_{ip} - \sum_{i\in\mathcal{V}_f}\sum_{j\in\mathcal{N}(i_f)\cap\mathcal{N}(i)}(v_i - v_j)^{\mathrm{T}}u_j\\ &-\sum_{j=1}^{\omega}\sum_{i\in F(j)}(v_i - v_j)^{\mathrm{T}}u_j + \sum_{i\in\mathcal{V}_f}\sum_{j\in\mathcal{N}(i_f)\cap\mathcal{N}(i)}(\alpha_{ij} - \bar{\alpha})|v_i - v_j|_1\\ &+\sum_{j=1}^{\omega}\sum_{i\in F(j)}(v_i - v_j)\sum_{p\in F(j)\cap\mathcal{N}(i)}(\alpha_{ip} - \bar{\alpha})\mathrm{sgn}\left((v_i - v_j) - (v_p - v_j)\right). \end{aligned} \tag{28}$$

Since V_{ip} is symmetric with respect to x_{ip} and $x_{ip} = (x_i - \chi) - (x_p - \chi)$, $\forall\chi \in \Re^m$, it holds

$$\begin{aligned}&\frac{1}{2}\sum_{i\in\mathcal{V}_f}\sum_{p\in\mathcal{V}_f\cap\mathcal{N}(i)} s_{\mathcal{L}(i)}\dot{x}_{ip}^{\mathrm{T}}\nabla_{x_{ip}}V_{ip}(x_{ip})\\ &= \frac{1}{2}\sum_{j=1}^{\omega}\sum_{i\in F(j)}\sum_{p\in F(j)\cap\mathcal{N}(i)}((v_i - v_j) - (v_p - v_j))\nabla_{x_{ji}}\hat{V}_{ip}((x_i - x_j) - (x_p - x_j))\\ &= \frac{1}{2}\sum_{j=1}^{\omega}\sum_{i\in F(j)}\sum_{p\in F(j)\cap\mathcal{N}(i)}\left((v_i - v_j)\nabla_{x_{ip}}\hat{V}_{ip}(x_{ip}) + (v_p - v_j)\nabla_{x_{pi}}\hat{V}_{pi}(x_{ip})\right)\\ &= \sum_{j=1}^{\omega}\sum_{i\in F(j)}(v_i - v_j)\sum_{p\in F(j)\cap\mathcal{N}(i)}\nabla_{x_{ip}}\hat{V}_i(x_{ip}). \end{aligned} \tag{29}$$

Since $u_j(t)$ is continuous for $t \in [0, \infty)$ and it is proved in Part A that $\lim_{t\to\infty} u_j = 0$, it holds that $u_j(t)$ is bounded for $t \in [0, \infty)$. Denote the bound as μ such that $|u_j(t)|_1 \leqslant \mu, \forall j \in \mathcal{N}(i_f)$. Substituting (29) into (28) yields

$$\begin{aligned}\dot{\Upsilon} \leqslant&\dot{H} + \sum_{i\in\mathcal{V}_f}\sum_{j\in\mathcal{N}(i_f)\cap\mathcal{N}(i)}|u_j||v_i - v_j| + \sum_{j=1}^{\omega}\sum_{i\in F(j)}|u_j||v_i - v_j| - \sum_{i\in\mathcal{V}_f}\sum_{j\in\mathcal{N}(i_f)\cap\mathcal{N}(i)}\bar{\alpha}|v_i - v_j|_1\\ &-\sum_{j=1}^{\omega}\sum_{i\in F(j)}(v_i - v_j)^{\mathrm{T}}\sum_{p\in F(j)\cap\mathcal{N}(i)}\bar{\alpha}\mathrm{sgn}((v_i - v_j) - (v_p - v_j)).\end{aligned}$$

Define the number of agents in $F(j)$, $j \in \{1, \ldots, \omega\}$ as $s_{F(j)}$. Define $\check{V}_j$ as a column stack vector of $(v_i - v_j)$, $i \in F(j)$. Let $\mathcal{G}_j$ be the undirected graph characterizing the interactions among the $s_{F(j)}$ followers of leader j with the associated Laplacian matrix $L_j \triangleq D_j D_j^{\mathrm{T}}$. Note that by definition, L_j is symmetric positive semi-definite. Let $\bar{\mathcal{G}}_j$ be the directed graph characterizing the interactions among leader j and its followers. Let the edge weight $a_{ij} = 1$ if leader j is a neighbor of follower i and $a_{ij} = 0$ otherwise. Define $\Lambda_j \triangleq \mathrm{diag}(a_{i_1 j}, \ldots, a_{i_{s_{F(j)}} j})$, $i_k \in F(j)$, $k = 1, \ldots, s_{F(j)}$. Note that $\Lambda_j^2 = \Lambda_j$ because $a_{ij}, i \in F(j)$ is either 1 or 0. Therefore, it holds

$$\dot{\Upsilon} \leqslant \dot{H} + \sum_{j=1}^{\omega}\mathbf{1}_{s_{F(j)}} \otimes \mu|\check{v}_j| - \sum_{j=1}^{\omega}\bar{\alpha}|\Lambda_j \otimes I_m\check{v}_j|_1 - \sum_{j=1}^{\omega}\bar{\alpha}|D_j^{\mathrm{T}} \otimes I_m\check{v}_j|_1.$$

The leader-follower topology matrix associated with graph $\bar{\mathcal{G}}_j$ is defined as $R_j \triangleq L_j + \Lambda_j$, $j \in \{1, \ldots, \omega\}$. According to [22], R_j is symmetric positive definite. Based on (26) and the fact that $|\cdot| \leqslant |\cdot|_1$ for any vector, one obtains

$$\dot{\Upsilon} \leqslant -\kappa_v\sum_{j\in\mathcal{N}(i_f)}(v_j - v_{i_f})^{\mathrm{T}}(v_j - v_{i_f}) - \frac{\kappa_v}{2}\sum_{j\in\mathcal{N}(i_f)}\sum_{p\in\mathcal{N}(i_f)\cap\mathcal{N}(j)}(v_j - v_p)^{\mathrm{T}}(v_j - v_p)$$

$$-\tilde{k}^{\mathrm{T}}(C_{i_f}^F)^{\mathrm{T}}C_{i_f}^F\tilde{k}-\sum_{j=1}^{\omega}\left(\bar{\alpha}\sqrt{\lambda_{\min}(R_j)}-\bar{\mu}\right)|\breve{v}_j|, \tag{30}$$

where $\bar{\mu}\triangleq\max_{k\in\{s_{F(1)},s_{F(2)},\ldots,s_{F(\omega)}\}}\{\mathbf{1}_k\otimes\mu\}$. If $R_j(t)$, changes at some time, there exists $t_1>0$ such that $R_j(t)=R_j(0)$, $\forall t\in[0,t_1)$. By designing $\bar{\alpha}>\bar{\mu}/\min_{j\in\{1,\ldots,\omega\}}\{\sqrt{\lambda_{\min}(R(j)(0))}\}$, one has $\dot{\Upsilon}(t)\leqslant 0$ for $t\in[0,t_1)$. Since $V_{ik}(t)$, $i\in\mathcal{V}'$, $k\in\mathcal{N}(j)\cap\mathcal{V}'$ is continuous, we can conclude that $V_{ik}(t_1)\leqslant\Upsilon(t_1)$. From the definition of V_{ik} in (2), it follows that there is no collision and no edge in the graph $\bar{\mathcal{G}}_j(0)$ will be lost for $t\in[0,t_1]$. Therefore, the only possibility for $R_j(t)$ to change at $t=t_1$ is if some edges are added in the graph, which means that $\bar{\mathcal{G}}_j(0)$ is a subgraph of $\bar{\mathcal{G}}_j(t_1)$. Then it yields from Lemma 2.2 in [23] that $R_j(0)\leqslant R_j(t_1),\forall j\in\{1,\ldots,\omega\}$ and thus $\lambda_{\min}(R_j(0))\leqslant\lambda_{\min}(R_j(t_1))$. Following the same argument, if $R_j(t)$ changes at $t=t_i>t_1$, $i=1,\ldots$, the same conclusion can be drawn. Therefore, it holds that $\bar{\mu}/\min_{j\in\{1,\ldots,\omega\}}\{\sqrt{\lambda_{\min}(R(j)(0))}\}\geqslant\bar{\mu}/\min_{j\in\{1,\ldots,\omega\}}\{\sqrt{\lambda_{\min}(R(j)(t))}\}$ for all $t\geqslant 0$. And there is no collision and also no edge in the graph $\bar{\mathcal{G}}_j(0)$ is lost for all $t\geqslant 0$. Thus, $\dot{\Upsilon}(t)\leqslant 0$, $\forall t\geqslant 0$.

According to the above analysis, it holds that $\lim_{t\to\infty}(v_j-v_{i_f})=0$ and $\lim_{t\to\infty}(v_i-v_j)=0$, $\forall i\in F(j),j\in\{1,\ldots,\omega\}$. The assumption that the initial undirected graph $\mathcal{G}'$ is connected indicates that there is at least one leader in $\{1,\ldots,\omega\}$ for each agent in $\mathcal{V}_f$ when $t=0$. Together with the fact that no edge in $\bar{\mathcal{G}}_j$, $\forall j\in\{1,\ldots,\omega\}$ is lost for $t\geqslant 0$, all agents in $\mathcal{V}_f$ are followers of agents in $\{1,\ldots,\omega\}$ for $t\geqslant 0$. This further leads to $\lim_{t\to\infty}(v_i-v_j)=0,\forall i,j\in\mathcal{V}$. This completes the proof.

Remark 2. For the case $s_a=1$, no desired configuration can be obtained by Lemma 3. Actually, even if $s_a=1$, a desired configuration may be found. If there exists $x^*_{i_fj},\forall j\in\mathcal{N}(i_f)$ satisfying $0<|x^*_{i_fj}|<R$ such that

$$\sum_{j\in\mathcal{N}(i_f)-a(i_f)}\nabla_{x_{i_f}}V_{i_fj}(|x^*_{i_fj}|)+\sum_{j\in a(i_f)}\nabla_{x_{i_f}}\tilde{V}_{i_fj}(|x^*_{i_fj}|)=0, \tag{31}$$

then $x_{ij}=x^*_{i_fj}$ is the desired displacement. The existence of the solution of (31) means that the forces acted on the malicious agent from its targets and non-target neighbors can be balanced together.

4 Flocking control against multiple malicious agents

In this section, multiple malicious agents in the swarm (1) are taken into consideration. Define the set of the malicious agents as $F\triangleq\{i_1,i_2,\ldots,i_d\}\subset\mathcal{V}$. Let $\mathcal{G}_\ell$ be the undirected graph characterizing the interactions among the agents in F. In the following, two cases are considered: Case 1, any two malicious agents are unconnected in $\mathcal{G}_\ell$; Case 2, any two malicious agents are connected in $\mathcal{G}_\ell$. The obtained results can be straightly extended to more general cases with both connected and unconnected malicious agents in $\mathcal{G}_\ell$.

4.1 Flocking control against multiple malicious agents in Case 1

Consider the swarm (1) with d ($d<N$) malicious agents, any two of which are unconnected in $\mathcal{G}_\ell$. The controller of the malicious agent $i_k\in F$, $k\in\{1,\ldots,d\}$ is

$$u_{i_k}=-k_v\sum_{j_k\in\mathcal{N}(i_k)}(v_{i_k}-v_{j_k})-\sum_{j_k\in\mathcal{N}(i_k)}\nabla_{x_{i_k}}\tilde{V}_{i_kj_k}, \tag{32}$$

where k_v and $\tilde{V}_{i_kj_k}$ are defined in (5). Define $C_{i_k}k\triangleq\sum_{j_k\in\mathcal{N}(i_k)}\nabla_{x_{i_k}}\tilde{V}_{i_kj_k}$ where $k\triangleq(k_v,k_a,k_r)^{\mathrm{T}}$ and $C_{i_k}\triangleq(\sum_{j_k\in\mathcal{N}(i_k)}(v_{i_k}-v_{j_k}),\sum_{j_k\in\mathcal{N}(i_k)}\nabla_{x_{i_k}}V_{ai_kj_k},\sum_{j_k\in\mathcal{N}(i_k)}\nabla_{x_{i_k}}V_{ri_kj_k})$.

Define an undirected graph $\mathcal{G}'_2\triangleq(\mathcal{V}'_2,\mathcal{E}'_2)$ consisting of a set of vertices $\mathcal{V}'_2\triangleq\mathcal{V}-F$ and the set of edges $\mathcal{E}'_2\triangleq\{(i,j)|i,j\in\mathcal{V}'_2,|x_{ij}|<R,i\neq j\}$.

Assumption 5. The initial graph $\mathcal{G}'_2$ is connected.

The flocking problem to be solved is reformulated as follows.

Problem $\mathcal{F}_2$. Consider the swarm (1) satisfying Assumption 5 with multiple malicious agents in F under controller (32) and any two of them are unconnected in $\mathcal{G}_\ell$. Design u_i, $i\in\mathcal{V}'_2$ such that

(1) $\lim_{t\to\infty}(v_i-v_j)=0,\ \forall i,j\in\mathcal{V}$;

(2) The swarm (1) asymptotically converges to a fixed geometric configuration, under which

• $u_{i_k}=0$ and $|x_{i_kj}|=\bar{\delta}_{i_kj}$, $\forall i_k\in F$, $j\in\mathcal{N}(i_k)$ where $0<\bar{\delta}_{ij}<R$;
• $|x_{ij}|=\tilde{\delta}_{ij}$, $\forall i\in\mathcal{V}_2'$, $j\in\mathcal{N}(i)$ where $0<\tilde{\delta}_{ij}<R$;
(3) $|x_{ij}(t)|\neq 0$ for $t\geqslant 0$, $\forall i,j\in\mathcal{V}$ and $i\neq j$.

Consequently, two scenarios are considered: one where malicious agents have common neighbors and another where they do not.

4.1.1 *Malicious agents with no common neighbors*

The desired geometric configuration formed by each malicious agent and its neighbors can be obtained separately. Therefore, the method in Lemma 1 can still be utilized here to identify the desired configurations for malicious agents in set F. Geometrically, there are d regular polygons with each malicious agent being the center and its neighbors being vertexes.

Similarly to the case when there is only one malicious agent, the following assumption is made to ensure the existence of the desired configuration.

Assumption 6. At the initial time, each malicious agent has at least two neighbors.

To deal with the unknown parameter k_{i_k} of the malicious agent $i_k\in F$, $k\in\{1,\ldots,d\}$, filtering functions in (32) for agent i_k, it holds that

$$\begin{aligned}&\dot{v}_{i_k}^F=-av_{i_k}^F+av_{i_k},\quad v_{i_k}^F(0)=v_{i_k}(0),\\&\dot{C}_{i_k}^F=-aC_{i_k}^F+C_{i_k},\quad C_{i_k}^F(0)=0,\end{aligned}$$

where $a>0$. $v_{i_k}^F$ and $C_{i_k}^F$ are the filtered functions of v_{i_k} and C_{i_k}, respectively.

Define $\hat{k}_{i_k}$ as the estimate of parameter k for agent i_k. Design another adaptive update law of $\hat{k}$ as follows:

$$\dot{\hat{k}}_{i_k}=-\Gamma_kC_{i_k}\sum_{j_k\in\mathcal{N}(i_k)}(v_{j_k}-v_{i_k})-\Gamma_k(C_{i_k}^F)^{\mathrm{T}}(C_{i_k}^F\hat{k}_{i_k}-v_{i_k}^F+v_{i_k}).\tag{33}$$

Define $\mathcal{V}_F\triangleq\{i_1,j_{11},j_{12},\ldots,i_2,j_{21},j_{22},\ldots,i_d,j_{d1},j_{d2},\ldots\}$ where $j_{kg}\in\mathcal{N}(i_k)$, $i_k\in F$, $k\in\{1,2,\ldots,d\}$, $g\in\{1,2,\ldots,\hat{s}_k\}$ with $\hat{s}_k>0$ being the number of agents in $\mathcal{N}(i_k)$. An improved controller for agent $j_k\in\mathcal{N}(i_k)$ is designed as follows:

$$u_{j_k}=-\kappa_v\sum_{p\in\mathcal{N}(j_k)\cap\mathcal{V}_F}(v_{j_k}-v_p)-\kappa_v(v_{j_k}-v_a)-\kappa_x\sum_{p\in\mathcal{N}(j_k)\cap\mathcal{V}_F}\nabla_{x_{j_k}}\hat{V}_{j_kp}-C_{i_k}\hat{k}_{i_k},\tag{34}$$

where $\tilde{V}_{i_kq}$ and $\hat{V}_{j_kp}$ are defined in (5) and (8), respectively. Velocity v_a satisfies

$$\sum_{k=1}^{d}\hat{s}_kv_a^2-\left(\sum_{i_k\in F}v_{i_k}^{\mathrm{T}}+\sum_{i_k\in F}\sum_{j_k\in\mathcal{N}(i_k)}v_{j_k}^{\mathrm{T}}\right)v_a+\sum_{i_k\in F}\sum_{j_k\in\mathcal{N}(i_k)}v_{i_k}^{\mathrm{T}}v_{j_k}\leqslant 0.\tag{35}$$

The hierarchical flocking control architecture introduced in Section 3 is also utilized here. Similarly to Figure 2, all malicious agents in F can be set in Layer 1, their normal neighbors in $\mathcal{V}_F-F$ are in Layer 2, and other agents in $\mathcal{V}-\mathcal{V}_F$ are in Layer 3. The controller (34) and (35) for agents in Layer 2 only utilizes information from Layers 1 and 2, with agents in Layers 2 and 3 acting as leaders and followers. Then, the following theorem is proposed.

Theorem 2. Consider the swarm (1) satisfying Assumptions 1, 5, and 6 with multiple malicious agents in F under controller (32), where any two of them are unconnected in $\mathcal{G}_\ell$. Suppose that the malicious agents have no common neighbors. Problem $\mathcal{F}_2$ can be solved by applying controller (34) and (35) along with parameter estimate update law (33) to agents in $\mathcal{V}_F-F$ and controller (10) to agents in $\mathcal{V}-\mathcal{V}_F$.

Proof. Define x_F and v_F as the vectors consisting of positions and velocities of all the agents in $\mathcal{V}_F$, respectively. Define $\tilde{k}_{i_k}\triangleq k_{i_k}-\hat{k}_{i_k}$. Define the energy function $G(x_F,v_F)$ as

$$\begin{aligned}G(x_F,v_F)\triangleq&\kappa_x\sum_{i_k\in F}\sum_{j_k\in\mathcal{N}(i_k)}\hat{V}_{j_ki_k}+\frac{\kappa_x}{2}\sum_{i_k\in F}\sum_{j_k\in\mathcal{N}(i_k)}\sum_{r_k\in\mathcal{V}_F\cap\mathcal{N}(j_k)-\{i_k\}}\hat{V}_{j_kr_k}\\&+\frac{1}{2}\sum_{i_k\in F}\sum_{j_k\in\mathcal{N}(i_k)}(v_{j_k}-v_{i_k})^{\mathrm{T}}(v_{j_k}-v_{i_k})+\frac{1}{2}\sum_{i_k\in F}\tilde{k}_{i_k}^{\mathrm{T}}\Gamma_k^{-1}\tilde{k}_{i_k}.\end{aligned}$$

By applying controller (34) and (35) along with estimate update law (33), the derivative of G holds that

$$\begin{aligned}\dot{G}(x_F,v_F)=&-\kappa_v\sum_{i_k\in F}\sum_{j_k\in\mathcal{N}(i_k)}\left((v_{j_k}-v_{i_k})^{\mathrm{T}}(v_{j_k}-v_{i_k})+\frac{1}{2}\sum_{p_k\in\mathcal{N}(j_k)\cap\mathcal{V}_F}(v_{j_k}-v_{p_k})^{\mathrm{T}}(v_{j_k}-v_{p_k})\right)\\&-k_v\sum_{i_k\in F}\sum_{j_k\in\mathcal{N}(i_k)}(v_{j_k}-v_{i_k})^{\mathrm{T}}\sum_{i_k\in F}\sum_{j_k\in\mathcal{N}(i_k)}(v_{j_k}-v_{i_k})-\sum_{i_k\in F}\sum_{j_k\in\mathcal{N}(i_k)}(v_{j_k}-v_{i_k})^{\mathrm{T}}(v_{j_k}-v_a)\\\leqslant&-\kappa_v\sum_{i_k\in F}\sum_{j_k\in\mathcal{N}(i_k)}(v_{j_k}-v_{i_k})^{\mathrm{T}}(v_{j_k}-v_{i_k})-\frac{\kappa_v}{2}\sum_{i_k\in F}\sum_{j_k\in\mathcal{N}(i_k)}\sum_{p_k\in\mathcal{N}(j_k)\cap\mathcal{V}_F}(v_{j_k}-v_{p_k})^{\mathrm{T}}(v_{j_k}-v_{p_k})\\&-\sum_{i_k\in F}\sum_{j_k\in\mathcal{N}(i_k)}(v_{j_k}-v_a)^{\mathrm{T}}(v_{j_k}-v_a)-\sum_{i_k\in F}\tilde{k}_{i_k}^{\mathrm{T}}(C_{i_k}^F)^{\mathrm{T}}C_{i_k}^F\tilde{k}_{i_k}.\end{aligned}\tag{36}$$

Therefore, $\dot{G}(x_F,v_F)$ is always non-positive. Define the level set $\Omega_F\triangleq\{(x_F^{\mathrm{T}},v_F^{\mathrm{T}})^{\mathrm{T}}\in R^{2s_F\times m}:G(t)\leqslant G(0),G(0)>0\}$ where $s_F\triangleq\sum_{k\in\{1,\ldots,d\}}(\hat{s}_k+1)$ for $t\geqslant 0$. Hence, by applying LaSalle's invariance principle, $(x_F^{\mathrm{T}},v_F^{\mathrm{T}})$ starting in Ω_F asymptotically converges to the largest invariant set inside the region $\mathcal{C}_{\mathcal{F}}\triangleq\{(x_F^{\mathrm{T}},v_F^{\mathrm{T}})^{\mathrm{T}}\in R^{2s_F\times m}:\dot{G}(t)=0\}$. According to (36), $\dot{G}(t)=0$ holds if and only if $v_{i_k}=v_{j_k}=v_a$ and $C_{i_k}^F\tilde{k}_{i_k}=0$ for all $i_k\in F$ and $j_k\in\mathcal{N}(i_k)$. This implies that $\lim_{t\to\infty}(v_i-v_j)=0$, $\forall i,j\in\mathcal{V}_F$. Under this result, by the similar analysis in the proof of Theorem 1, one concludes that the flocking Problem $\mathcal{F}_2$ is solved.

Remark 3. Comparing the improved controller (34) and (35) with controller (8) in Theorem 1, the term $-\kappa_v(v_{j_k}-v_a)$ for $j_k\in\mathcal{N}(i_k)$, $i_k\in F$ in (34) is added to guarantee that all agents in $\mathcal{V}_F$ can tend to the common velocity v_a. Displacement and velocity information can be obtained either through a centralized information interaction and control architecture or by building a local communication network among agents in $\mathcal{V}_F$ if they are close to each other. Conversely, if they are far apart such that the velocity information of all agents in $\mathcal{V}_F$ cannot be obtained by agent j_k, then by building a local communication network among agents in $\mathcal{N}(j_k)$, $\forall j_k\in\mathcal{N}_F$, only the "local" flocking goal can be guaranteed, that is, each malicious agent and its neighbors tend to a common velocity, respectively.

4.1.2 *Malicious agents with common neighbors*

Here, the common neighbors are affected by more than one malicious agent. To contain the malicious agents, a specific geometric configuration needs to be designed.

For agent $j\in\mathcal{V}_F-F$, find all the malicious agents that are neighbors of agent j and define the number of them as $\bar{s}_j$. Denote the set of them as $\mathcal{N}(j)\cap F\triangleq\{i_{j1},i_{j2},i_{j3},\ldots\}$ where $i_{jk}\in F$ and $k\in\{1,2,\ldots,\bar{s}_j\}$. Solve the following equations:

$$\begin{cases}\sum_{l1\in\mathcal{N}(i_{j1})-F}\nabla_{x_{i_{j1}l1}^*}\tilde{V}_{i_{j1}l1}=0,\\\sum_{l2\in\mathcal{N}(i_{j2})-F}\nabla_{x_{i_{j2}l2}^*}\tilde{V}_{i_{j2}l2}=0,\\\qquad\vdots\\\sum_{l\bar{s}_j\in\mathcal{N}(i_{j\bar{s}_j})-F}\nabla_{x_{i_{j2}l\bar{s}_j}^*}\tilde{V}_{i_{j\bar{s}_j}l\bar{s}_j}=0,\\0<|x_{i_{jk}j}^*|<R,\quad k=1,2,\ldots,\bar{s}_j,\end{cases}\tag{37}$$

the solutions provide the desired displacement between the malicious agents and their neighbors. In this case, the total potential gradient of all the malicious agents with respect to their neighbors is restricted to 0 by designing all their neighbors' relative positions simultaneously. Physically, this means that the forces acting on each malicious agent by all its neighbors are balanced, ensuring that each malicious agent remains contained in the swarm. Figure 3 illustrates a desired configuration satisfying (37). The malicious agents i_{f_1} and i_{f_2} which are unconnected in $\mathcal{G}_\ell$ have a common neighbor j_2. Agents j_1-j_5 are their neighbors. Under this configuration, agents i_{f_1} and i_{f_2} are contained simultaneously.

Define $\mathcal{N}_F\triangleq\mathcal{V}_F-F\triangleq\{1,2,\ldots,s_2\}$ as the set of normal agents that are neighbors of one or multiple malicious agents with s_2 representing the number of these neighbors. Define $\mathcal{N}_m\triangleq\{1,2,\ldots,s_1\}$ as the set of normal agents that are neighbors of multiple malicious agents and $\mathcal{N}_o\triangleq\{s_1+1,s_1+2,\ldots,s_2\}$ as the set of normal agents that are neighbors of only one malicious agent where $s_2>s_1\geqslant 1$. Design the

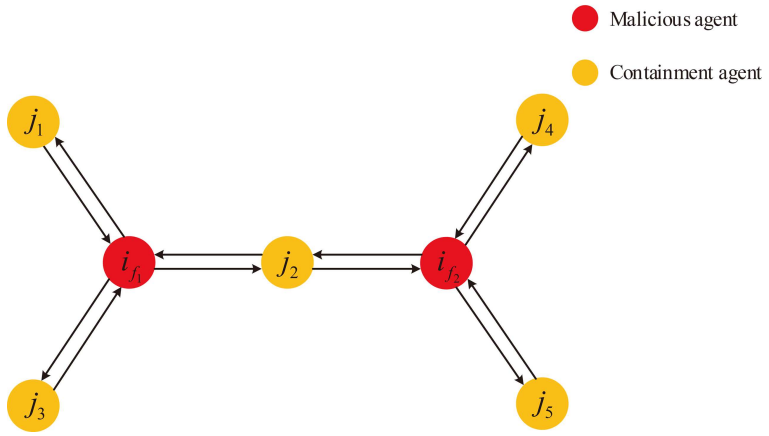

Figure 3 (Color online) Illustration of the configuration to contain malicious agents which are unconnected in $\mathcal{G}_\ell$ with a common neighbor.

controller for agent $j \in \mathcal{N}_F$ as follows:

$$u_j = -\kappa_v \sum_{i_p \in \mathcal{N}(j)\cap F} (v_j - v_{i_p}) - \kappa_v (v_j - v_b) - \kappa_x \sum_{p \in \mathcal{N}(j)\cap \mathcal{V}_F} \nabla_{x_j} \hat{V}_{jp} - \sum_{i_k \in \mathcal{N}(j)\cap F} C_{i_k} \hat{k}_{i_k}, \tag{38}$$

where $\hat{V}_{jp}$ is defined in (8), $\tilde{V}_{i_k q}$ is defined in (5) and $\hat{k}_{i_k}$ is defined in (33). v_b satisfies

$$\begin{aligned} &\left(\sum_{j=1}^{s_1} \bar{s}_j + s_2 - s_1\right) v_b^2 - \left(\sum_{j=1}^{s_1} \sum_{i_p \in \mathcal{N}(j)\cap F} (v_j + v_{i_p})^{\mathrm{T}} + \sum_{j=s_1+1}^{s_2} (v_j + v_{i_j})^{\mathrm{T}}\right) v_b \\ &+ \sum_{j=1}^{s_1} \sum_{i_p \in \mathcal{N}(j)\cap F} v_{i_p}^{\mathrm{T}} v_j + \sum_{s_1+1}^{s_2} v_{i_j}^{\mathrm{T}} v_j \leqslant 0, \end{aligned} \tag{39}$$

where $\bar{s}_j$ is defined in (37). The malicious agent i_j is the neighbor of agent $j \in \mathcal{N}_o$.

Similarly to the analysis in Theorem 2, the following result can be obtained.

Corollary 1. Consider swarm (1) satisfying Assumptions 1 and 5 with multiple malicious agents in F under controller (32) and any two of them are unconnected in $\mathcal{G}_\ell$. If there exist solutions in (37) for agents in $\mathcal{N}_F$, Problem $\mathcal{F}_2$ can be solved by applying controller (38) along with parameter estimate update law (33) to agents in $\mathcal{N}_F$ and controller (10) to agents in $\mathcal{V} - \mathcal{V}_F$.

4.2 Flocking control against multiple malicious agents in Case 2

Considering there are d malicious agents in swarm (1) and any two of them are connected in $\mathcal{G}_\ell$, the controller of the malicious agent i_k, $k \in \{1, \ldots, d\}$, $i_k \in F$ is

$$u_{i_k} = -\sum_{i_g \in F - \{i_k\}} (v_{i_k} - v_{i_g}) - \sum_{i_p \in \mathcal{N}(i_k)\cap F} \nabla_{x_{i_k}} V_{i_k i_p} - \sum_{j_k \in \mathcal{N}(i_k) - F} \nabla_{x_{i_k}} \tilde{V}_{i_k j_k} - \sum_{i_g \in F - \{i_k\}} \sum_{j_g \in \mathcal{N}(i_g) - F} \nabla_{x_{i_g}} \tilde{V}_{i_g j_g}, \tag{40}$$

where the malicious potential $\tilde{V}_{i_k j_k}$ is defined in (5). Define $C_2 k_2 \triangleq \sum_{i_k \in F} \sum_{j_k \in \mathcal{N}(i_k) - F} \nabla_{x_{i_k}} \tilde{V}_{i_k j_k}$, where $C_2 \triangleq \left(\sum_{i_k \in F} \sum_{j_k \in \mathcal{N}(i_k) - F} \nabla_{x_{i_k}} V_{a i_k j_k}, \sum_{i_k \in F} \sum_{j_k \in \mathcal{N}(i_k) - F} \nabla_{x_{i_k}} V_{a i_k j_k}\right)$ and $k_2 \triangleq (k_a, k_r)^{\mathrm{T}}$. Define $f_2 \triangleq \sum_{i_g \in F - \{i_k\}} (v_{i_k} - v_{i_g}) + \sum_{i_p \in \mathcal{N}(i_k)\cap F} \nabla_{x_{i_k}} V_{i_k i_p}$.

Remark 4. These multiple malicious agents are regarded as "accomplices", and they cooperatively damage the swarm, as reflected in the last term of (40). For any malicious agent, only its normal neighbors are treated as "adversaries". Thus, in the controller (40), the malicious behavior only targets normal agents rather than the malicious ones.

For swarm (1) with the above malicious agents, the flocking problem is reformulated as follows.

Problem $\mathcal{F}_3$. Consider the swarm (1) satisfying Assumptions 5 with multiple malicious agents in F under controller (40) and any two of them are connected in $\mathcal{G}_\ell$. Design u_i, $i \in \mathcal{V}_2'$ such that the goals (1)–(3) in Problem $\mathcal{F}_2$ can be achieved.

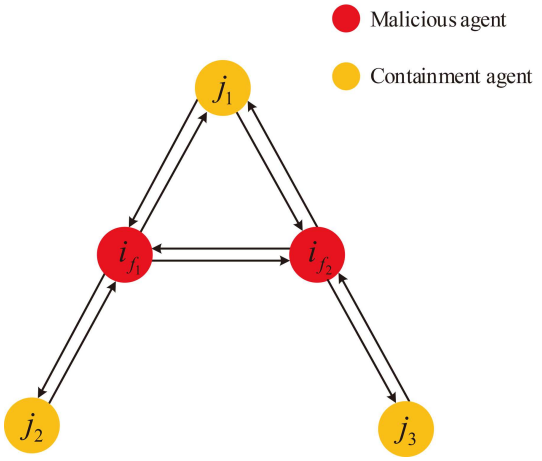

Figure 4 (Color online) Illustration of the configuration to contain malicious agents which are connected in $\mathcal{G}_\ell$.

Since each normal agent may suffer from the combined effects caused by multiple malicious agents that are connected in $\mathcal{G}_\ell$, the desired configuration to contain these malicious agents can also be obtained by solving (37). Figure 4 illustrates an example of the desired configuration satisfying (37) with two malicious agents i_{f1} and i_{f2} that are connected in $\mathcal{G}_\ell$. Agents j_1–j_3 are their neighbors. Under this geometric configuration, the behaviors of agents i_{f1} and i_{f2} are contained simultaneously.

To deal with the unknown control parameter k_2 in (40), let

$$\begin{aligned}&\sum_{i_k\in F}\dot{v}_{i_k}^F=-a\sum_{i_k\in F}v_{i_k}^F+a\sum_{i_k\in F}v_{i_k},\quad \sum_{i_k\in F}v_{i_k}^F(0)=\sum_{i_k\in F}v_{i_k}(0),\\&\dot{f}_2^F=-af_2^F+f_2,\quad f_2^F(0)=0,\\&\dot{C}_2^F=-aC_2^F+C_2,\quad C_2^F(0)=0,\end{aligned}$$

where $a>0$. $\sum_{i_k\in F}v_{i_k}^F$, f_2^F and C_2^F are the filtered functions of $\sum_{i_k\in F}v_{i_k}$, f_2 and C_2.

Define $\hat{k}_2$ as the estimate of k_2 and design its adaptive update law as

$$\dot{\hat{k}}_2=-\Gamma_{k_2}C_2^{\mathrm{T}}\sum_{j\in\mathcal{N}_F}\sum_{i_j\in F\cap\mathcal{N}(j)}(v_j-v_{i_j})-\Gamma_{k_2}(C_2^F)^{\mathrm{T}}\left(C_2^F\hat{k}_2-\sum_{i_k\in F}v_{i_k}^F+\sum_{i_k\in F}v_{i_k}+f_2^F\right),\tag{41}$$

where $\Gamma_{k_2}>0$ is the gain matrix.

With the desired configuration obtained by (37) and the estimate (41), the geometric configuration controller is designed for agent $j\in\mathcal{N}_F$ as

$$\begin{aligned}u_j=&-\sum_{i_j\in\mathcal{N}(j)\cap F}\left((\kappa_v(v_j-v_{i_j})+\kappa_x\nabla_{x_j}\hat{V}_{ji_j})+\sum_{i_h\in\mathcal{N}(i_j)\cap F}\nabla_{x_{i_j}}V_{i_ji_h}\right)\\&-\sum_{i_j\in\mathcal{N}(j)\cap F}\sum_{i_g\in F-\{i_j\}}(v_{i_j}-v_{i_g})-C_2\hat{k}_2,\end{aligned}\tag{42}$$

where κ_v, κ_x and $\hat{V}_{ji_j}$ are defined in (8). $V_{i_ji_h}$ is defined in (3) and C_2 is defined in (40).

Theorem 3. Consider the swarm (1) satisfying Assumptions 1 and 5 with multiple malicious agents in F under controller (40) and any two of them are connected in $\mathcal{G}_\ell$. If solutions to (37) exist for agents in $\mathcal{N}_F$, Problem $\mathcal{F}_3$ can be solved by applying controller (42) along with the parameter estimate update law (41) to agents in $\mathcal{N}_F$ and controller (10) to agents in $\mathcal{V}-\mathcal{V}_F$.

Proof. Define $\tilde{k}_2\triangleq k_2-\hat{k}_2$. Let $d>0$ be the number of malicious agents. Define the energy function of all the agents in $\mathcal{V}_F$ as

$$Y(x_F,v_F)\triangleq\kappa_x\sum_{j\in\mathcal{N}_F}\sum_{i_j\in F\cap\mathcal{N}(j)}\hat{V}_{ji_j}+(d+2)\sum_{i_k\in F}\sum_{i_p\in\mathcal{N}(i_k)\cap F}V_{i_ki_p}+\sum_{i_k\in F}\sum_{i_g\in F-\{i_k\}}(v_{i_k}-v_{i_g})^{\mathrm{T}}(v_{i_k}-v_{i_g})$$

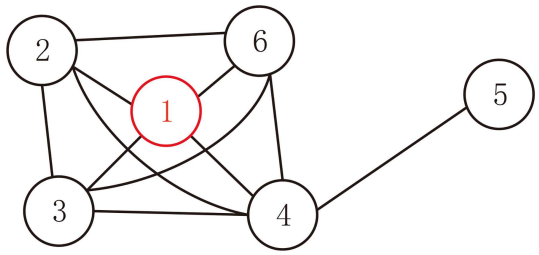

Figure 5 (Color online) Illustration of the initial topology of agents in Example 1.

$$+\frac{1}{2}\sum_{j\in\mathcal{N}_F}\left(\sum_{i_j\in F\cap\mathcal{N}(j)}(v_j-v_{i_j})^{\mathrm{T}}\right)\left(\sum_{i_j\in F\cap\mathcal{N}(j)}(v_j-v_{i_j})\right)+\frac{1}{2}\tilde{k}_2^{\mathrm{T}}\Gamma_{k_2}^{-1}\tilde{k}_2.$$

Combining controllers (10), (40), (42), and parameter estimate update law (41), the time derivative of Y is

$$\begin{aligned}\dot{Y}=&(d+2)\sum_{i_k\in F}\sum_{i_p\in\mathcal{N}(i_k)\cap F}(v_{i_k}-v_{i_p})^{\mathrm{T}}\nabla_{x_{i_k}}V_{i_ki_p}\Bigg(-2(v_{i_k}-v_{i_g})\\&-\sum_{i_e\in F-\{i_k,i_g\}}(v_{i_k}-v_{i_e})+\sum_{i_e\in F-\{i_k,i_g\}}(v_{i_g}-v_{i_e})\Bigg)-\sum_{i_k\in F}\sum_{i_p\in\mathcal{N}(i_k)\cap F}\Bigg(\sum_{i_r\in F-\{i_k,i_p\}}(v_{i_k}-v_{i_r})\nabla_{x_{i_k}}V_{i_ki_p}\\&+\sum_{i_c\in F-\{i_k,i_p\}}(v_{i_p}-v_{i_c})\nabla_{x_{i_p}}V_{i_pi_k}+2(v_{i_k}-v_{i_p})\nabla_{x_{i_k}}V_{i_ki_p}\Bigg)\\&-\sum_{j\in\mathcal{N}_F}\Bigg(\sum_{i_j\in F\cap\mathcal{N}(j)}(v_j-v_{i_j})^{\mathrm{T}}\Bigg)\Bigg(\sum_{i_j\in F\cap\mathcal{N}(j)}\left(\kappa_v(v_j-v_{i_j})+\kappa_x\nabla_{x_j}\hat{V}_{ji_j}\right)\Bigg)-\tilde{k}_2^{\mathrm{T}}(C_2^F)^{\mathrm{T}}C_2^F\tilde{k}_2\\=&(d+2)\sum_{i_k\in F}\sum_{i_p\in\mathcal{N}(i_k)\cap F}(v_{i_k}-v_{i_p})^{\mathrm{T}}\nabla_{x_{i_k}}V_{i_ki_p}-\sum_{i_k\in F}d(v_{i_k}-v_{i_g})^{\mathrm{T}}(v_{i_k}-v_{i_g})\\&-\sum_{i_k\in F}\sum_{i_p\in\mathcal{N}(i_k)\cap F}\Bigg(\sum_{i_r\in F-\{i_k\}}(v_{i_k}-v_{i_r})-\sum_{i_c\in F-\{i_p\}}(v_{i_p}-v_{i_c})+2(v_{i_k}-v_{i_p})\Bigg)\nabla_{x_{i_k}}V_{i_ki_p}\\&-\kappa_v\sum_{j\in\mathcal{N}_F}\Bigg(\sum_{i_j\in F\cap\mathcal{N}(j)}(v_j-v_{i_j})^{\mathrm{T}}\Bigg)\Bigg(\sum_{i_j\in F\cap\mathcal{N}(j)}(v_j-v_{i_j})\Bigg)-\tilde{k}_2^{\mathrm{T}}(C_2^F)^{\mathrm{T}}C_2^F\tilde{k}_2\\=&-d\sum_{i_k\in F}(v_{i_k}-v_{i_g})^{\mathrm{T}}(v_{i_k}-v_{i_g})-\kappa_v\sum_{j\in\mathcal{N}_F}\Bigg(\sum_{i_j\in F\cap\mathcal{N}(j)}(v_j-v_{i_j})^{\mathrm{T}}\Bigg)\Bigg(\sum_{i_j\in F\cap\mathcal{N}(j)}(v_j-v_{i_j})\Bigg)\\&-\tilde{k}_2^{\mathrm{T}}(C_2^F)^{\mathrm{T}}C_2^F\tilde{k}_2,\end{aligned}$$

which means that $\dot{Y}\leqslant 0$. Similarly to the proof of Theorem 1, one can deduce that Problem $\mathcal{F}_3$ is solved using controllers (10) and (42). This completes the proof.

Remark 5. When the multiple malicious agents share common neighbors, the desired configuration of $\mathcal{V}_F$ to contain the malicious behaviors is obtained by solving equation (37). Note that $0<|x^*_{i_jkj}|<R$ is a restrictive condition, under which (37) may have no solution. Then, no desired configuration can be found, and the multiple malicious agents cannot reach a fixed common velocity.

One may refer to [24] for the leader-follower flocking control with time-varying velocities by regarding the malicious agents as leaders in the swarm.

5 Simulation

In this section, several numerical examples are presented to illustrate the effectiveness of proposed flocking control schemes in the above sections.

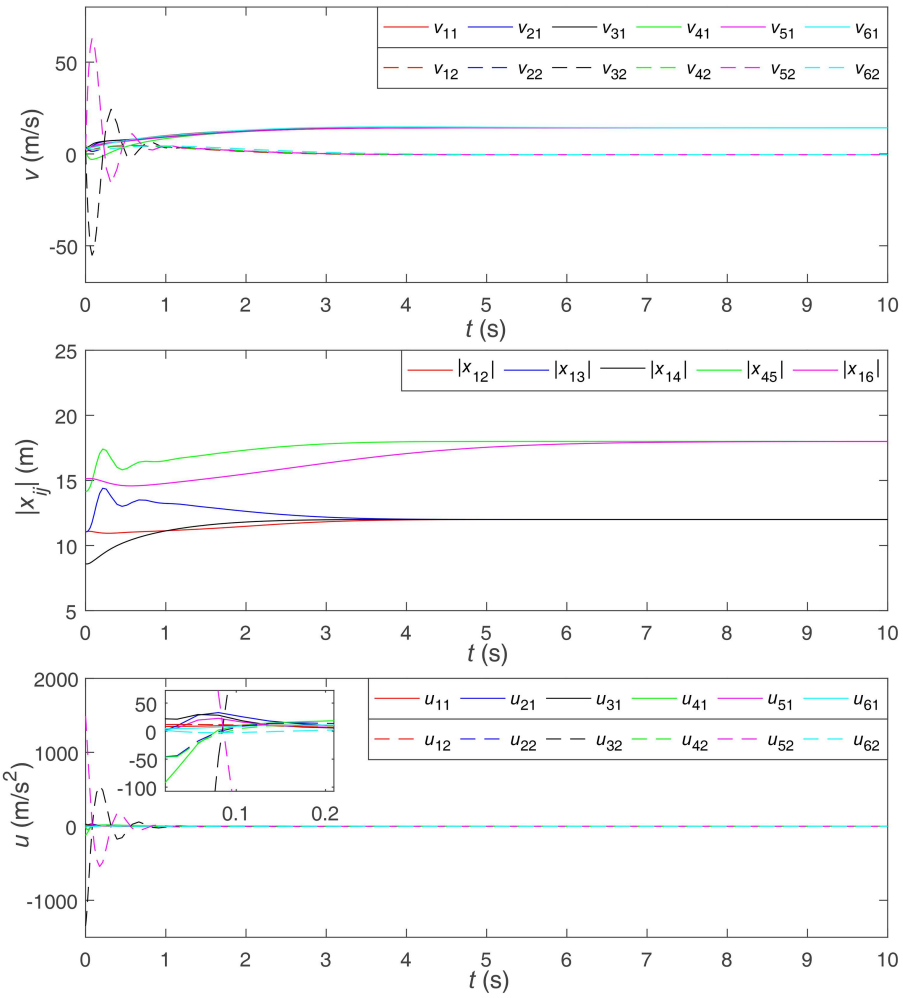

Figure 6 (Color online) Trajectories of velocities, distances between neighboring agents and control efforts of Agents 1–6 in Example 1.

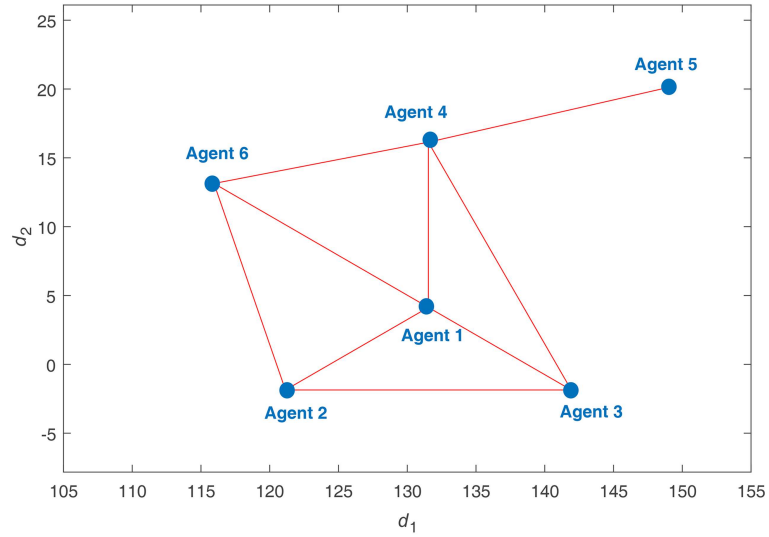

Figure 7 (Color online) Flocking patterns at $t = 10$ s where d_1 and d_2 present positions in the first and second dimensions in Example 1.

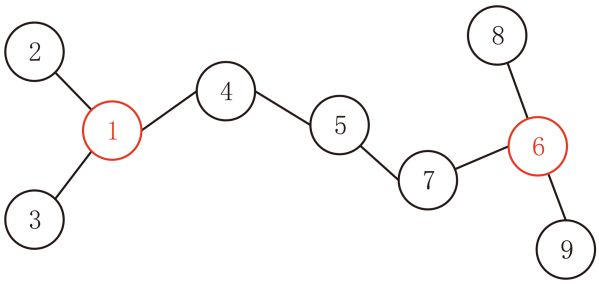

Figure 8 (Color online) Illustration of the initial topology of agents in Example 2.

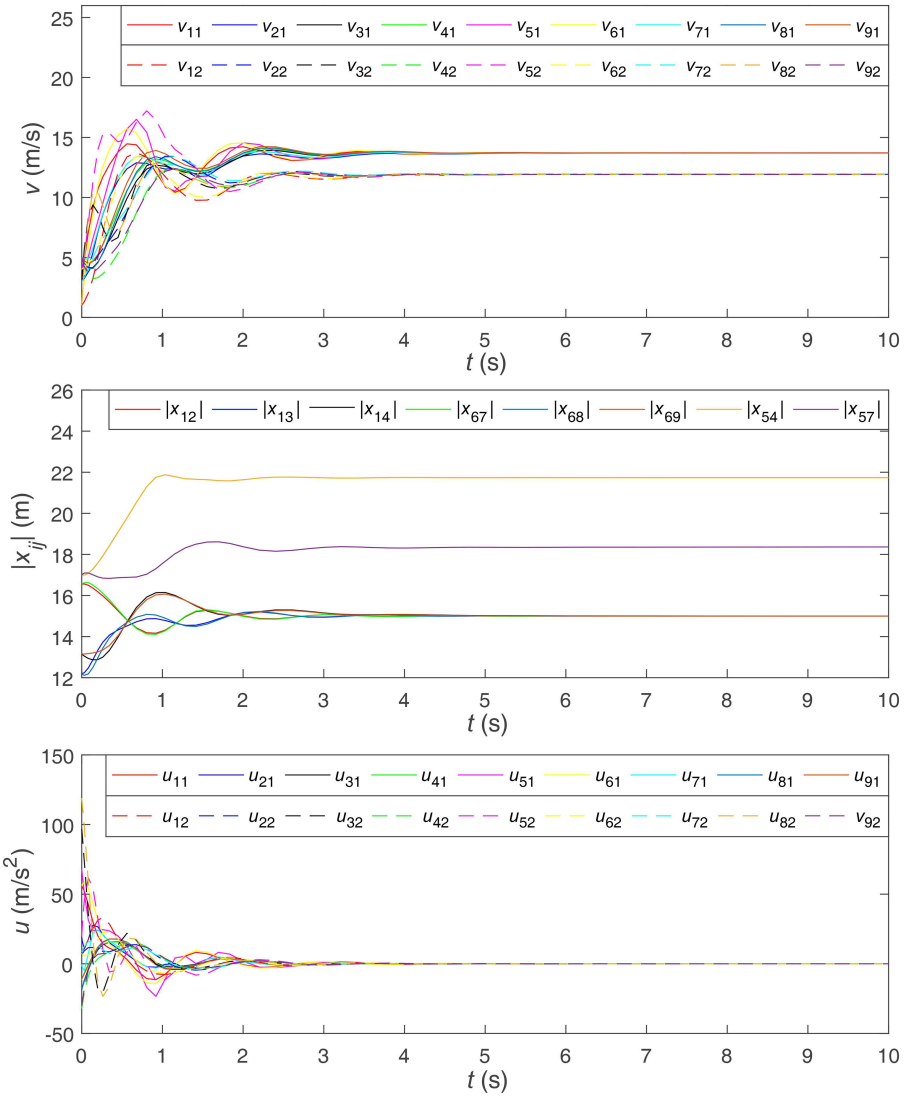

Figure 9 (Color online) Trajectories of velocities, distances between neighboring agents, and control efforts of Agents 1–9 in Example 2.

5.1 Example 1: one malicious agent with targets

Consider a 2-dimensional swarm of 6 agents (Agents 1–6), including a malicious agent (Agent 1) under controller (4) and (5) with $k_a = 0$, $k_r = 10^6$ and $k_v = 0$. The targets of it are Agents 2–4. Define the velocity of agent $i \in \mathcal{V}$ as $v_i \triangleq (v_{i1}, v_{i2})$ where v_{i1} and v_{i2} are velocities in the first and second dimensions, respectively. The initial undirected topology graph is shown in Figure 5. The initial values of velocities are randomly taken from $(0, 6)$ m/s $\times$ $(0, 6)$ m/s. Let the communication distance be $R = 18\sqrt{2}$ m. Let

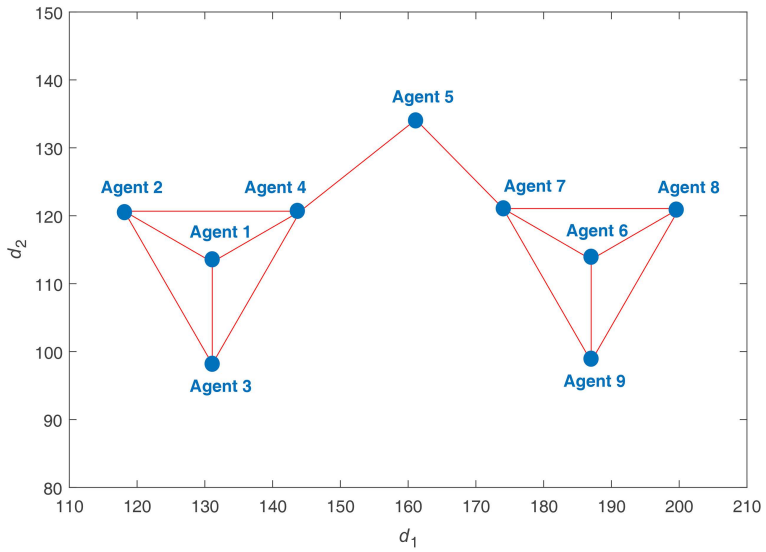

Figure 10 (Color online) Flocking patterns at $t = 10$ s where d_1 and d_2 present positions in the first and second dimensions in Example 2.

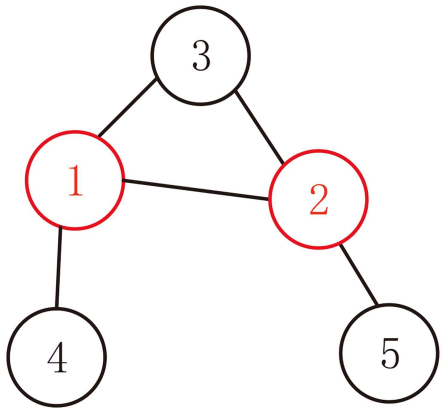

Figure 11 (Color online) Illustration of the initial topology of agents in Example 3.

the desired distance between the malicious agent and its neighbors be $\bar{\delta} = 12$ m. Apply controller (8) with $\kappa_v = 6.5$ and $\kappa_x = 20$ to Agents 2–4 and 6, and controller (10) with $\gamma_{kp} = 2$ to Agent 5. One can see from the simulation results in Figure 6 that all agents tend to a common velocity and all the control efforts tend to 0. Agent 1 is contained, and the distances between it and its targets approach 12 m. The distance between Agent 1 and its non-target neighbor, Agent 6, approaches 18 m. The flocking patterns at $t = 10$ s are presented in Figure 7.

5.2 Example 2: two malicious agents which are unconnected in $\mathcal{G}_\ell$

Consider a 2-dimensional swarm of 9 agents (Agents 1–9), including two malicious agents which are unconnected in $\mathcal{G}_\ell$ (Agents 1 and 6) under controller (32) with $k_a = 0, k_r = 10^3, k_v = 1$. The initial undirected topology graph is shown in Figure 8. Define the velocity of agent $i \in \mathcal{V}$ as $v_i \triangleq (v_{i1}, v_{i2})$ where v_{i1} and v_{i2} are velocities in the first and second dimensions, respectively. The initial values of velocities are randomly taken from $(0, 6)$ m/s $\times$ $(0, 6)$ m/s. Let the communication distance be $R = 18\sqrt{2}$ m. Let the desired distance between the malicious agent and its neighbors be $\bar{\delta} = 15$ m. Apply controller (34) and (35) with $\kappa_v = 2$ and $\kappa_x = 12$ to Agents 2–4 and 7–9, and controller (10) with $\gamma_{kp} = 3$ to Agent 5. The simulation results in Figure 9 illustrate that all agents tend to a common velocity. Malicious Agents 1 and 6 are contained, and the distances between them and their neighbors approach 15 m. Agent 5 is not a neighbor of any malicious agent and the distances between it and its neighbors tend to be fixed. The configuration of the agents at $t = 10$ s is shown in Figure 10.

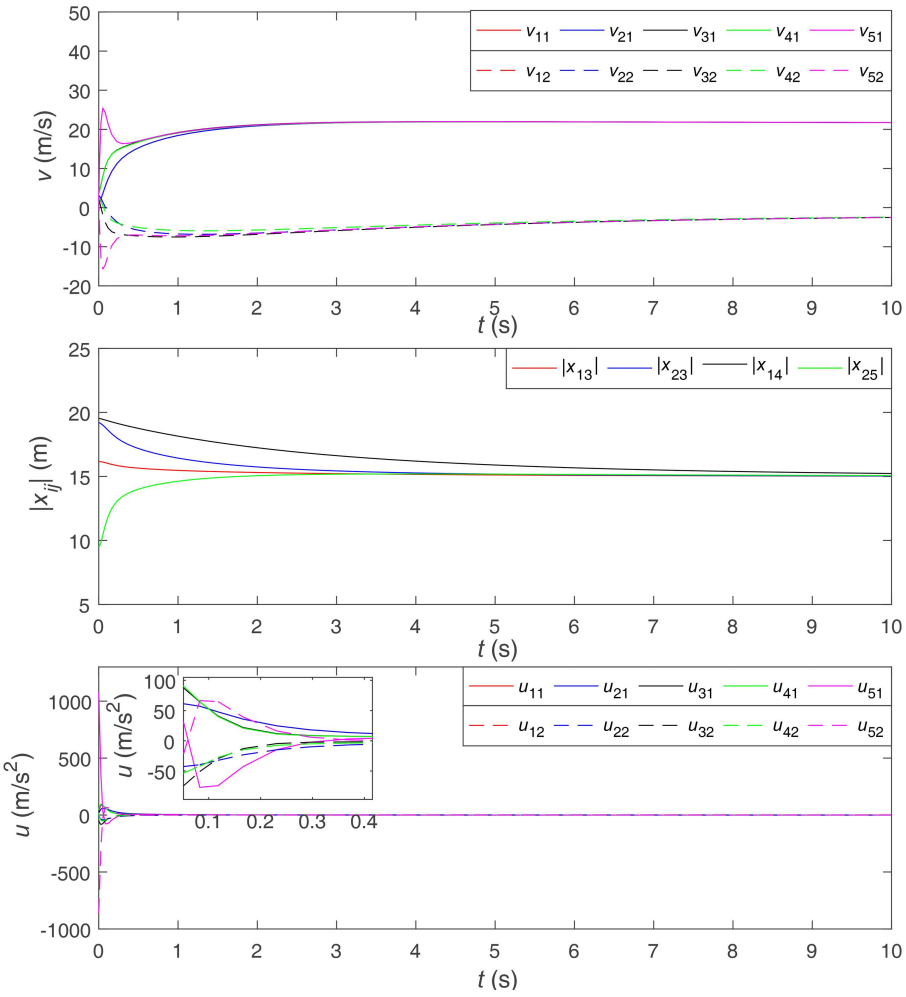

Figure 12 (Color online) Trajectories of velocities and distances between neighboring agents and control efforts of Agents 1–5 in Example 3.

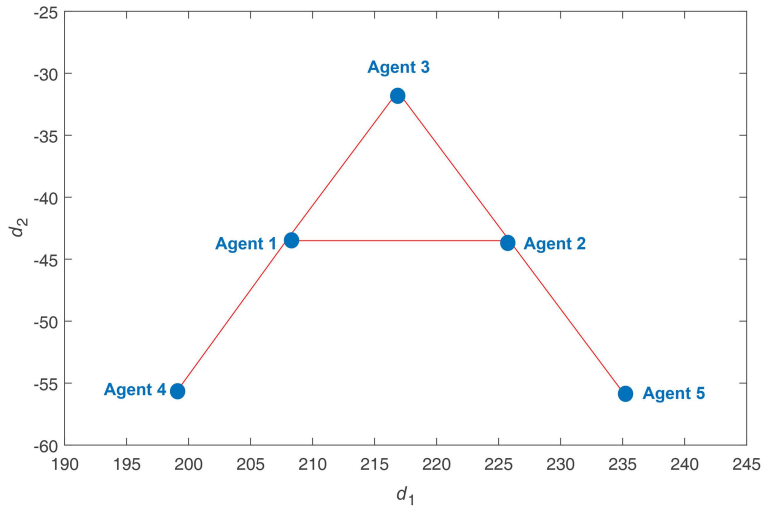

Figure 13 (Color online) Flocking patterns at $t = 10$ s where d_1 and d_2 present positions in the first and second dimensions in Example 3.

5.3 Example 3: two malicious agents which are connected in $\mathcal{G}_\ell$

Consider a 2-dimensional swarm of 5 agents (Agents 1–5), including two malicious agents which are connected in $\mathcal{G}_\ell$ (Agents 1 and 2) under controller (40) with $k_a = 0, k_r = 10^5, k_v = 0$. The initial undirected topology graph is shown in Figure 11. Define the velocity of agent $i \in \mathcal{V}$ as $v_i \triangleq (v_{i1}, v_{i2})$ where v_{i1} and v_{i2} are velocities in the first and second dimensions, respectively. The initial values of velocities are randomly taken from $(0, 6)$ m/s $\times$ $(0, 6)$ m/s. Let the communication distance be $R = 25$ m and the desired distance between the malicious agent and its neighbors be 18 m. Apply controller (42) with $\kappa_v = 4$ and $\kappa_x = 40$ to Agents 3–5. From the simulation results in Figure 12, one can see that all agents tend to have a common velocity. The malicious Agents 1 and 2 are contained simultaneously. The distances among them and their neighbors tend to 18 m as expected and the configuration tends to be the desired one, as shown in Figure 13.

6 Conclusion

This paper proposed hierarchical geometric configuration-based methods to design fault-tolerant flocking control for a swarm with malicious agents. The geometric containment conditions ensure that each malicious agent exerts zero force by appropriately designing the relative positions of each malicious agent's neighbors. Based on these conditions, a specific potential function was constructed, aiding in the design of the controllers for each malicious agent's neighbors. The methods were applicable to several common practical scenarios: single malicious agent with targets, multiple congregate malicious agents, and scattered ones. These new results addressed the problem without removing the malicious agents, thereby enriching the fault-tolerant flocking control theory. However, our work still has some limitations. When containing multiple clustered malicious agents, the displacements between malicious agents and all their neighbors need to be adjusted together. In addition, the malicious agents we considered were based on a specific model, and other types of malicious agents with different forms have not been considered. Further investigation is required to deal with the flocking control problem involving multiple malicious agents, including both scattered and clustered ones simultaneously. More complicated or general malicious behaviors could be investigated with the help of cognitive behavioral theory [25], game theory [26], and other related decision-making theories [27]. In addition, the application of containment methods to multi-agent systems with complex nonlinear dynamics presents an interesting topic for future research.

Acknowledgements This work was supported in part by National Natural Science Foundation of China (Grant Nos. 62073165, 62233009) and Funding of Postgraduate Research & Practice Innovation Program of Jiangsu Province (Grant No. KYCX21-0222).

References

1 Chen Y, Zhu Q, Xu H. Finding rough set reducts with fish swarm algorithm. Knowl-Based Syst, 2015, 81: 22–29
2 Varol Altay E, Alatas B. Bird swarm algorithms with chaotic mapping. Artif Intell Rev, 2020, 53: 1373–1414
3 Olfati-Saber R. Flocking for multi-agent dynamic systems: algorithms and theory. IEEE Trans Automat Contr, 2006, 51: 401–420
4 Tanner H G, Jadbabaie A, Pappas G J. Flocking in fixed and switching networks. IEEE Trans Automat Contr, 2007, 52: 863–868
5 Li Y, Li Y X, Tong S. Event-based finite-time control for nonlinear multiagent systems with asymptotic tracking. IEEE Trans Automat Contr, 2023, 68: 3790–3797
6 Reynolds C W. Flocks, herds and schools: a distributed behavioral model. In: Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques, 1987. 25–34
7 Fernando M. Online flocking control of UAVs with mean-field approximation. In: Proceedings of IEEE International Conference on Robotics and Automation (ICRA), 2021. 8977–8983
8 Beaver L E, Malikopoulos A A. An overview on optimal flocking. Annu Rev Control, 2021, 51: 88–99
9 Amin A A, Hasan K M. A review of fault tolerant control systems: advancements and applications. Measurement, 2019, 143: 58–68
10 Li Y, Zhao Y, Liu W, et al. Adaptive fuzzy predefined-time control for third-order heterogeneous vehicular platoon systems with dead zone. IEEE Trans Ind Inf, 2023, 19: 9525–9534
11 Gutiérrez León P, García-Morales J, Escobar-Jiménez R F, et al. Implementation of a fault tolerant system for the internal combustion engine's MAF sensor. Measurement, 2018, 122: 91–99
12 Ahmed A, Abu Bakar K, Channa M I, et al. A survey on trust based detection and isolation of malicious nodes in ad-hoc and sensor networks. Front Comput Sci, 2015, 9: 280–296
13 Raya M, Papadimitratos P, Aad I, et al. Eviction of misbehaving and faulty nodes in vehicular networks. IEEE J Sel Areas Commun, 2007, 25: 1557–1568
14 Hobbs K L, Cargal C, Feron E, et al. Early safety analysis of manned-unmanned team system. In: Proceedings of AIAA Information Systems-AIAA Infotech @ Aerospace, 2018
15 Zhang C, Yang H, Jiang B, et al. Flocking control against the malicious agent. 2023. ArXiv:2308.04285
16 Li S, Guo J, Xiu J, et al. Attacking cooperative multi-agent reinforcement learning by adversarial minority influence. 2023. ArXiv:2302.03322

17 Kang M S, Lee S B, Gligor V D. The crossfire attack. In: Proceedings of IEEE Symposium on Security and Privacy, 2013. 127–141
18 Wang W, Li H, Sun Y, et al. CatchIt: detect malicious nodes in collaborative spectrum sensing. In: Proceedings of IEEE Global Telecommunications Conference, 2009. 1–6
19 Shang Y. Consensus of hybrid multi-agent systems with malicious nodes. IEEE Trans Circ Syst II, 2020, 67: 685–689
20 Dong Y, Huang J. Flocking with connectivity preservation of multiple double integrator systems subject to external disturbances by a distributed control law. Automatica, 2015, 55: 197–203
21 Hendrickson B. Conditions for unique graph realizations. SIAM J Comput, 1992, 21: 65–84
22 Mei J, Ren W, Ma G. Distributed coordinated tracking for multiple Euler-Lagrange systems. In: Proceedings of the 49th IEEE Conference on Decision and Control, 2010. 3208–3213
23 Ghapani S, Mei J, Ren W, et al. Fully distributed flocking with a moving leader for Lagrange networks with parametric uncertainties. Automatica, 2016, 67: 67–76
24 Yu W, Chen G, Cao M. Distributed leader-follower flocking control for multi-agent dynamical systems with time-varying velocities. Syst Control Lett, 2010, 59: 543–552
25 Kanellopoulos A, Vamvoudakis K G. Non-equilibrium dynamic games and cyber-physical security: a cognitive hierarchy approach. Syst Control Lett, 2019, 125: 59–66
26 Tembine H, Zhu Q, Basar T. Risk-sensitive mean-field games. IEEE Trans Automat Contr, 2013, 59: 835–850
27 Cao M. Merging game theory and control theory in the era of AI and autonomy. Natl Sci Rev, 2020, 7: 1122–1124