A Leader-Follower Attack-Tolerant Algorithm for Resilient Rendezvous with Reduced Network Redundancy

Zirui Liao, *Graduate Student Member, IEEE,* Jian Shi, Yuwei Zhang, Shaoping Wang, Rentong Chen, and Zhiyong Sun, *Member, IEEE*

***Abstract*—This paper addresses the resilient rendezvous problem for leader-follower multi-agent systems (MASs) in the presence of adversarial attacks. A novel leader-follower attack-tolerant (LFAT) algorithm is developed to ensure that the healthy followers reach rendezvous on the reference value propagated by healthy leaders. Compared with the existing weighted mean-subsequence-reduced (W-MSR) algorithm, the proposed LFAT algorithm includes a necessary state initialization step for leader and follower agents and an improved threat elimination step, so that more effective information can be retained for state updates. The necessary and sufficient condition on the network topology is further derived to ensure resilient rendezvous for leader-follower MASs. Compared with the existing resilient algorithms, the proposed LFAT algorithm enables MASs to achieve leader-follower resilient rendezvous under relaxed graph robustness conditions, so that the network redundancy is mitigated. Several numerical examples are given to illustrate the superior performance of the LFAT algorithm and the scalability to lager-scale and time-varying networks.**

***Index Terms*—Leader-follower resilient rendezvous, multi-agent system, LFAT algorithm, adversarial attack, network topology**

I. Introduction

With the development of modern industrial automation, guaranteeing and improving the resilience of the multi-agent system (MAS) against adversarial attacks has become increasingly critical [1]–[5] due to its wide application in multi-robot systems [6], smart grids [7], wireless sensor networks [8], and formation control [9], [10]. The lack of global situational awareness and distributed characteristics render MASs vulnerable to targeted adversarial attacks. Agents may be compromised by a cyber attack, which would render the agent lose its capability to implement the prescribed control protocols. More seriously, a single misbehaving agent injection can cause an overall system collapse [11]. Consequently, it is vital to investigate resilient coordination problems for MASs, so that healthy agents can operate normally under the influence of adversarial attacks.

Toward this objective, the *resilient rendezvous problem* has been extensively investigated in recent years [12]–[16]. Such problems can be roughly divided into the leaderless resilient rendezvous problem [12] and the leader-follower resilient rendezvous problem [13]. The former aims to devise distributed resilient algorithms to ensure that the state values of healthy agents reach rendezvous on a certain value or an error range, regardless of the influence of some misbehaving agents. The latter usually requires a leader-follower structure and resilient algorithms to ensure that the follower agents (followers) reach rendezvous on the reference value propagated by a single or multiple leader agents (leaders) [17], [18]. Followers should not only track the leaders' reference state, but also defend against the influence of misbehaving agents, who may send abnormal information to other healthy agents. Compared to the leaderless resilient rendezvous problem, addressing the leader-follower resilient rendezvous problem is more challenging, since both leaders and followers may be compromised by attackers and more stringent graph conditions may be necessitated [11].

In the context of resilient rendezvous, a class of algorithms known as *Mean-Subsequence-Reduced* (MSR) has garnered increasing attention in the past decade owing to its attack-tolerant feature. In the seminal study [12], the authors developed a weighted MSR (W-MSR) method to ensure resilient rendezvous against Byzantine attacks. It helps each healthy agent eliminate potential contaminated information received from its in-neighbor set. Motivated by this pioneering work, numerous variants of the W-MSR algorithm have been developed [19]–[24]. Under leader-follower structures, the paper [13] adopted the W-MSR algorithm to address the leader-follower resilient rendezvous problem for MASs with single-integrator dynamics. Subsequently, the leader-follower resilient rendezvous with dynamic leaders [14], [25] and over time-varying networks [26] were achieved based on the idea of MSR, respectively. In addition, the study [27] extended the resilient rendezvous results above to containment control and achieved resilient containment for leader-follower MASs with single-integrator and double-integrator dynamics, respectively. However, most of these MSR-based results merely present

This work was supported in part by the National Natural Science Foundation of China under Grant 62173147, Grant 62303030, Grant 62403028 and Grant U2233212; in part by the Beijing Municipal Natural Science Foundation under Grant L221008, and in part by the Open Fund of Science and Technology on Thermal Energy and Power Laboratory under Grant TPL2022C02. *(Corresponding author: Shaoping Wang.)*

Z. Liao, J. Shi, Y. Zhang, S. Wang and R. Chen are with the School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China.

S. Wang is with the Tianmushan Laboratory, Hangzhou 310023, China, and also with the Beihang-Ningbo Innovation Research Institute, Beihang University, Ningbo 315800, China. (e-mail: shaopingwang@buaa.edu.cn)

Z. Sun is with State Key Laboratory for Turbulence and Complex Systems, Department of Mechanics and Engineering Science, College of Engineering, Peking University, Beijing 100091, China.

Z. Liao is with the Shenyuan Honors College, Beihang University, Beijing 100191, China.

sufficient graph conditions to achieve leader-follower resilient rendezvous or containment. To enhance theoretical completeness, a further investigation of the necessary graph condition is required.

In addition to the lack of complete theoretical conditions, another concern for leader-follower resilient control is the high redundancy of the network topology. Although MSR-type algorithms can efficiently eliminate the potential contaminated data for healthy agents, they usually require that each healthy agent possesses a sufficient number of in-neighbors (i.e., graph robustness) to implement the threat elimination step. This requirement induces a redundant network topology. As a consequence, existing results [13], [26], [27] related to resilient control are built on stringent graph conditions, which require the communication topology to be strongly $(2f+1)$-robust or even strongly $(3f+1)$-robust, where f refers to the maximum number of misbehaving agents in the in-neighbor set of each healthy agent. More seriously, the network will become extremely complex as the parameter f increases, which would result in heavy topology redundancy.

To mitigate the stringent graph condition, the idea of the trusted node was introduced in [28] to ensure resilient rendezvous for MASs despite any number of adversarial attacks. Such node is assumed to be sufficiently secure and cannot be compromised by adversarial attacks. Motivated by this seminal study, trusted-nodes-based methods were developed in [13], [29], [30] to improve the graph robustness and achieve resilient rendezvous for MASs with reduced network redundancy. Similar resilient algorithms incorporating trusted edges [31] and trusted regions [32] were developed accordingly to relax the graph robustness requirements. Nevertheless, these trusted units are usually assumed to be free from adversarial attacks, which turns out to be a stringent assumption and may be difficult to satisfy in practical scenarios. For example, the paper [13] requires all leaders to be trusted to guarantee resilient rendezvous for leader-follower MASs with reduced network robustness. Thus, how to reduce network redundancy from an algorithmic perspective becomes a challenging problem.

Inspired by the aforementioned observations and motivated by the W-MSR algorithm, we develop a leader-follower attack-tolerant (LFAT) algorithm to guarantee resilient rendezvous for leader-follower MASs in the presence of adversarial attacks with reduced network redundancy. The research emphases of this study are summarized as follows.

1) We develop a novel LFAT algorithm to address the resilient rendezvous problem for leader-follower MASs. Compared to the studies [13], [26], [27], the proposed LFAT algorithm relaxes the sufficient graph condition for achieving leader-follower resilient rendezvous from strong $(2f+1)$-robustness of the whole network to strong $(f+1)$-robustness of the healthy network, which means that fewer directed edges are required and topology redundancy is thereby reduced. The construction of the proposed graph condition in practical scenarios has also been explicitly stated.
2) With the proposed LFAT algorithm, we further derive the necessary and sufficient graph condition to achieve leader-follower resilient rendezvous. In contrast to the existing results [14], [25], [33] on leader-follower resilient rendezvous that only partially addressed sufficient graph conditions, this study proves that strong $(f+1)$-robustness of the healthy network is the necessary and sufficient condition to achieve resilient rendezvous for leader-follower MASs, which bridges the gap on graph condition analysis.
3) We further extend the obtained results to larger-scale MASs, time-varying networks, and microgrid systems, respectively, where the reference value propagated by healthy leaders may jump with time steps. The numerical examples show that the proposed LFAT algorithm is still applicable to these complex scenarios. More specifically, healthy followers achieve resilient rendezvous despite the influence of multiple misbehaving agents, whether the reference value jumps or the network topology varies. Regarding the frequency synchronization problem in microgrid systems, the LFAT algorithm still ensures resilient synchronization of frequency reference corrections. The scalability of the proposed method is thus validated.

The remainder of this study is organized as follows. Section II introduces some preliminaries on graphs and formulates the resilient rendezvous problem for leader-follower MASs. The convergence analysis with the LFAT algorithm is conducted in Section III. Several numerical examples are given in Section IV to validate that strong $(f+1)$-robustness is necessary and sufficient to achieve resilient rendezvous for leader-follower MASs. The conclusion and future work are stated in Section V.

II. Problem Setup

A. Graph Theory Notions

Consider a leader-follower MAS modeled by a time-invariant digraph $\mathcal{D}=(\mathcal{V},\mathcal{E})$, where $\mathcal{V}=\{1,\ldots,n\}$ is the node set and $\mathcal{E}\subseteq\mathcal{V}\times\mathcal{V}$ is the directed edge set. The edge $(j,i)\in\mathcal{E}$ indicates that agent i has access to the information of agent j. The in-neighbor and out-neighbor sets for agent i are defined as $\mathcal{N}_i^{\text{in}}=\{j\in\mathcal{V}|(j,i)\in\mathcal{E}\}$ and $\mathcal{N}_i^{\text{out}}=\{j\in\mathcal{V}|(i,j)\in\mathcal{E}\}$, respectively. The inclusive in-neighbor set for agent i is further defined as $\mathcal{I}_i^{\text{in}}=\mathcal{N}_i^{\text{in}}\cup\{i\}$.

We consider the situation that the node set $\mathcal{V}$ contains misbehaving agents. Therefore, two essential notions for time-invariant graphs are given to characterize the resilience, i.e., reachability and robustness.

Definition 1 ([12]): Consider a time-invariant digraph $\mathcal{D}=(\mathcal{V},\mathcal{E})$ and a nonempty proper subset $\mathcal{S}\subset\mathcal{V}$. The set $\mathcal{S}$ is said to be r-reachable if it holds $\left|\mathcal{N}_i^{\text{in}}\setminus\mathcal{S}\right|\geq r,\ \exists\, i\in\mathcal{S}$, where $r\in\mathbb{Z}_+$.

Definition 2 ([34]): Consider a time-invariant digraph $\mathcal{D}=(\mathcal{V},\mathcal{E})$ and a nonempty proper subset $\mathcal{S}_1\subset\mathcal{V}$. The digraph $\mathcal{D}$ is said to be strongly r-robust with respect to (w.r.t.) $\mathcal{S}_1$ if for any nonempty subset $\mathcal{S}_2\subseteq\mathcal{V}\backslash\mathcal{S}_1$, $\mathcal{S}_2$ is r-reachable, where $r\in\mathbb{Z}_+$.

The concept of strongly r-robust graph requires that for any nonempty subset $\mathcal{S}_2\subseteq\mathcal{V}\backslash\mathcal{S}_1$, there exist some nodes in $\mathcal{S}_2$ who possess at least r in-neighbors from outside. This

property is crucial for achieving leader-follower rendezvous under adversarial attacks, which will be examined later.

The variants of reachable sets and robust graphs in time-varying networks are presented below, respectively.

Definition 3 (Jointly r-reachable set [35]): Consider a time-varying digraph $\mathcal{D}[t] = (\mathcal{V}, \mathcal{E}[t])$ and a nonempty proper subset $\mathcal{S} \subset \mathcal{V}$. $\mathcal{S}$ is said to be jointly r-reachable if there exists an infinite sequence of uniformly bounded and non-overlapping time intervals $[t_h, t_{h+1})$ such that in each time interval $[t_h, t_{h+1})$, there exist a time step $t_j \in [t_h, t_{h+1})$ and an agent $i_j \in \mathcal{S}$ such that $|\mathcal{N}_{i_j}^+[t_j] \backslash \mathcal{S}| \geq r$, where $r \in \mathbb{Z}_+$.

Definition 4 (Jointly and strongly r-robust graph [30]): Consider a time-varying digraph $\mathcal{D}[t] = (\mathcal{V}, \mathcal{E}[t])$ and a nonempty proper subset $\mathcal{S}_1 \subset \mathcal{V}$. $\mathcal{D}[t]$ is said to be jointly and strongly r-robust w.r.t. $\mathcal{S}_1$ if for any nonempty subset $\mathcal{S}_2 \subseteq \mathcal{V} \backslash \mathcal{S}_1$, $\mathcal{S}_2$ is jointly r-reachable, where $r \in \mathbb{Z}_+$.

B. Leader-Follower MAS

In this study, we consider a leader-follower MAS, which is composed of two types of agents: 1) leader agents (or leaders); and 2) follower agents (or followers), with the former propagating a reference value and the latter aiming to converge to this value and achieve resilient rendezvous. We denote the sets of leaders and followers as $\mathcal{L}$ and $\mathcal{F}$, respectively. For each leader $l \in \mathcal{L}$, its state remains constant:

$$x_l[t+1] = \Upsilon, \tag{1}$$

where $x_l[t+1] \in \mathbb{R}$ is the state value of leader l at time step $t+1$ with $t \in \mathbb{Z}_{\geq 0}$, and Υ remains constant over a period of time but may suddenly change at some time step. For each follower $i \in \mathcal{F}$, its state update follows

$$x_i[t+1] = F_i\left(\left\{x_j^i[t]\right\}\right), \tag{2}$$

where $x_j^i[t] \in \mathbb{R}$ is the value transmitted from agent $j \in \mathcal{I}_i^{\text{in}}$ to agent i at time step t with $x_i[t] = x_i^i[t]$, and the update rule $F_i(\cdot)$ is dependent on $x_j^i[t]$ and can be an arbitrary function. Regarding the leader-follower rendezvous problem, a typical control protocol [36] for followers is introduced as

$$F_i\left(\left\{x_j^i[t]\right\}\right) = \sum_{j \in \mathcal{I}_i^{\text{in}}} \theta_{ij}[t] x_j^i[t], \tag{3}$$

where $\theta_{ij}[t]$ is the weight of edge (j, i).

In the context of adversarial attack, agents in the leader-follower MAS are classified into healthy agents (or healthy nodes) and misbehaving agents (or misbehaving nodes) according to the following definitions:

Definition 5 ([12]): An agent is said to be healthy if it sends its current state $x_i[t]$ to all of its out-neighbors at each time step t and uses the rules (1) or (2) for state updates.

Definition 6 ([12]): An agent is said to be misbehaving if it sends its current state $x_i[t]$ to all of its out-neighbors at each time step t, but its state update is uncontrolled by the designed rule (manipulated by attackers).

Let the sets of healthy and misbehaving agents be $\mathcal{H}$ and $\mathcal{M}$, where it holds $\mathcal{H} \subseteq \mathcal{V}$ and $\mathcal{M} := \mathcal{V} \backslash \mathcal{H}$. Let $\mathcal{D}_h$ be the digraph of $\mathcal{H}$, which captures the healthy nodes and the connections between these nodes. Furthermore, let the sets of healthy leaders and healthy followers be $\mathcal{L}_h$ and $\mathcal{F}_h$. The following definition quantifies the distribution of adversarial attacks by limiting the maximum number of misbehaving agents in the in-neighbor set of each follower.

Definition 7 (f-local attack model [12]): The misbehaving set $\mathcal{M}$ is said to satisfy the f-local attack model if the in-neighbor set of each agent i contains at most f misbehaving agents, i.e., $\left|\mathcal{N}_i^{\text{in}} \cap \mathcal{M}\right| \leq f, \ \forall i \in \mathcal{V}$, where $f \in \mathbb{Z}_{\geq 0}$.

Remark 1: Since both leaders and followers may be compromised by adversarial attacks, the prescribed rules (1) and (2) will only be followed by healthy leaders and healthy followers for state updates. In addition, the identities of leaders and misbehaving agents are unknown to other agents in the network. This anonymity increases the difficulty of designing resilient algorithms.

Remark 2: The parameter f in Definition 7 does not have to be the exact number of misbehaving agents, but can be an upper limit. Furthermore, the f-local model is scalable, since it does not constrain the maximum number of misbehaving agents in the whole network. These properties show the applicability of the f-local model in practical scenarios. Therefore, the f-local model is adopted to characterize the distribution of adversarial attacks.

C. Resilient Rendezvous Problem

This study aims to tackle the following coordination problem. Specifically, we need to design resilient algorithms such that the state values of all healthy followers asymptotically reach agree on the reference value defined by the state values of healthy leaders, when the network is under the influence of misbehaving agents.

Definition 8 (Resilient rendezvous): A leader-follower MAS is said to achieve resilient rendezvous if

$$\lim_{t \to \infty} |x_i[t] - \Upsilon| = 0, \ \forall i \in \mathcal{F}_h \tag{4}$$

holds for all initial states of nodes and any possible misbehaving set.

Remark 3: The existing studies [12], [23], [24] mainly focus on addressing resilient rendezvous problems under leaderless structures, where the healthy agents are required to move within a convex hull (determined by the initial states of healthy agents) and reach rendezvous on a common state. However, it is also desirable that the healthy agents reach rendezvous on a certain reference state, which may be outside of this convex hull. This objective will be achieved by solving the above resilient rendezvous problem for leader-follower MASs.

D. LFAT Algorithm

Without prior information on the identity of misbehaving agents, how to eliminate potential threats for each healthy follower and achieve resilient rendezvous for the MAS remains a challenging problem. Motivated by [13], [37], a leader-follower attack-tolerant (LFAT) algorithm is developed in this subsection. The LFAT algorithm includes the following four steps:

1) *(Initializing states and parameters):* Initialize the state of leaders with Υ, initialize the state of followers randomly within a given interval, and set the attack-related parameter $f > 0$.
2) *(Collecting in-neighbors' information):* Each healthy follower i collects $x_i[t]$ and $x_j^i[t]$ from all $j \in \mathcal{N}_i^{\text{in}}$ at each time step t and sorts them in an ascending order.
3) *(Eliminating suspicious states):* Define the parameter $\varepsilon \in \mathbb{N}$ as

$$\varepsilon = \begin{cases} \max\left(|\mathcal{N}_i^{\text{in}}| - f - 1, 0\right), & \text{if } |\mathcal{N}_i^{\text{in}}| \leq 2f, \\ f, & \text{otherwise.} \end{cases} \tag{5}$$

Then, compared with $x_i[t]$, agent i eliminates the ε largest and smallest in-neighbor states in the sorted list. If there are fewer than ε in-neighbor states strictly larger or smaller than $x_i[t]$, then all of them will be eliminated.
4) *(Updating state values):* The healthy follower i applies

$$\begin{aligned} F_i\left(\{x_j^i[t]\}\right) &= \sum_{j \in \mathcal{I}_i^{\text{in}}[t] \backslash \mathcal{R}_i^{\text{in}}[t]} \theta_{ij}[t] x_j^i[t], \\ x_i[t+1] &= \theta_{ii}[t] x_i[t] + \sum_{j \in \mathcal{N}_i^{\text{in}}[t] \backslash \mathcal{R}_i^{\text{in}}[t]} \theta_{ij}[t] x_j^i[t] \end{aligned} \tag{6}$$

for state update, where $\theta_{ii}[t] = 1 - \sum_{j \in \mathcal{N}_i^{\text{in}}[t] \backslash \mathcal{R}_i^{\text{in}}[t]} \theta_{ij}[t]$ and $\mathcal{R}_i^{\text{in}}[t]$ is the set of removed in-neighbors through Step 3.

Note that Step 1 of the LFAT algorithm requires all leaders to have the same initial state, so that the healthy followers can converge to the identical value propagated by these leaders and achieve leader-follower rendezvous. Even if some healthy leaders or followers may be attacked subsequently and become misbehaving agents, their adverse effects will be eliminated through Steps 2 and 3.

Compared to the W-MSR algorithm, the LFAT algorithm not only explicitly states the initialization for agents' states and attack-related parameter, but also eliminates different in-neighbor values according to the number of in-neighbors. When $|\mathcal{N}_i^{\text{in}}| \leq 2f$, fewer or no in-neighbors will be removed. This operation is beneficial for healthy followers to approach the reference state under proper graph conditions. More importantly, with the LFAT algorithm, the sufficient condition for ensuring leader-follower resilient rendezvous can be relaxed from strongly $(2f+1)$-robustness to strongly $(f+1)$-robustness of $\mathcal{D}_h$, as we will see later.

Assumption 1: For all time steps $t \in \mathbb{N}$ and for all healthy agents $i \in \mathcal{H}$, the weights $\theta_{ii}[t]$ and $\theta_{ij}[t]$ in (6) satisfy the following conditions.

1) $\theta_{ij}[t] \geq \rho,\ \forall j \in \mathcal{I}_i^{\text{in}}$, where $\rho \in (0, 1)$;
2) $\theta_{ij}[t] = 0$ if $j \notin \mathcal{I}_i^{\text{in}}$;
3) $\sum_{j=1}^{n} \theta_{ij}[t] = 1$, where $n = |\mathcal{V}|$.

Remark 4: For Step 1 of the LFAT algorithm, it should be noted that initializing the leaders' states with Υ is a fixed and essential procedure. Otherwise, the resilient rendezvous problem will become a resilient containment problem [30], where the states of healthy followers converge to the interval constructed by the minimum and maximum values of the initial states of healthy leaders. In addition, since different random initialization strategies for followers may incur a minor increase in convergence time, it is advantageous to estimate an approximate initialization interval in advance to improve convergence performance.

Remark 5: The reason why the boundary is taken as "$2f$" in Eq. (5) is that the proposed leader-follower attack-tolerant (LFAT) algorithm eliminates different numbers of in-neighbor states (defined as ε) according to whether they have more than $2f$ in-neighbors or not. For one thing, this choice aligns with the W-MSR algorithm that eliminates up to $2f$ in-neighbors' states. For another, the boundary value "$2f$" not only guarantees resilient rendezvous, but also retains more effective information for state updates. Thus, resilient rendezvous can be achieved with reduced network redundancy. Furthermore, this boundary can better distinguish the two cases in Eq. (5).

III. Main Results

With the LFAT algorithm, the graph condition to achieve resilient rendezvous for leader-follower MASs is further derived in this section. In what follows, let

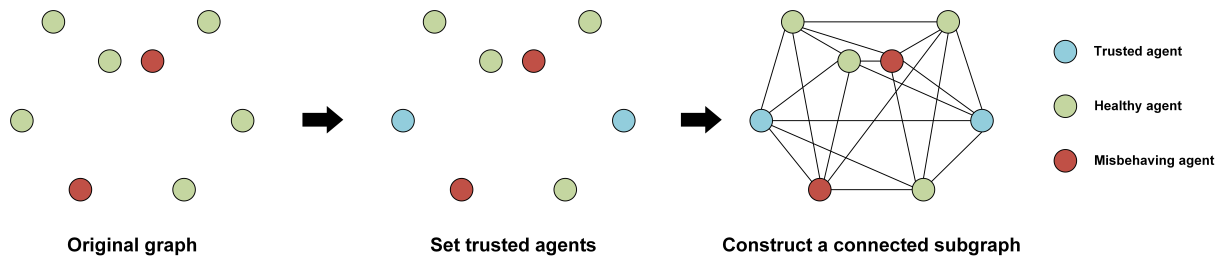
$$\begin{aligned} \overline{M}[t] &= \max_{i \in \mathcal{L}_h, j \in \mathcal{F}_h} \{x_i[t], x_j[t]\}, \\ \underline{m}[t] &= \min_{i \in \mathcal{L}_h, j \in \mathcal{F}_h} \{x_i[t], x_j[t]\}. \end{aligned} \tag{7}$$

Then, we present the following lemma for $x_j^i[t]$.

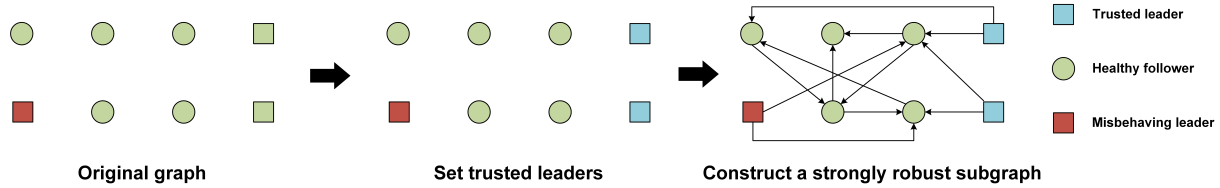
Lemma 1: Consider a leader-follower MAS modeled by $\mathcal{D} = (\mathcal{V}, \mathcal{E})$. Let Assumption 1 hold. Suppose that $\mathcal{D}_h$ is strongly $(f+1)$-robust w.r.t. $\mathcal{L}_h$, the misbehaving set $\mathcal{M}$ fulfills the f-local attack model, and the healthy followers execute the LFAT algorithm for state updates. Then, it holds $x_j^i[t] \in \left[\underline{m}[t], \overline{M}[t]\right],\ \forall i \in \mathcal{F}_h,\ \forall j \in \mathcal{N}_i^{\text{in}}[t] \backslash \mathcal{R}_i^{\text{in}}[t]$.

Proof: Since $\mathcal{D}_h$ is strongly $(f+1)$-robust w.r.t. $\mathcal{L}_h$ and the misbehaving set $\mathcal{M}$ fulfills the f-local attack model, for each healthy follower $i \in \mathcal{F}_h$, it possesses at least $f+1$ healthy in-neighbor agents and at most f misbehaving in-neighbor agents. Without loss of generality, we assume that Agent i possesses $f+1+a$ in-neighbors from $\mathcal{H}$ and b in-neighbors from $\mathcal{M}$, where $a, b \in \mathbb{N}$ and $b \leq f$. Consider $x_j^i[t] \notin \left[\underline{m}[t], \overline{M}[t]\right]$. Then, it must be sent by $\mathcal{M}$ according to the definitions of $\underline{m}[t]$ and $\overline{M}[t]$. When $|\mathcal{N}_i^{\text{in}}| \leq 2f$, it follows from (5) that $\varepsilon = a + b \geq b$. Through the LFAT algorithm, this misbehaving $x_j^i[t]$ will be eliminated. Otherwise, when $|\mathcal{N}_i^{\text{in}}| > 2f$, it yields $\varepsilon = f \geq b$. The misbehaving $x_j^i[t]$ will also be eliminated using the LFAT algorithm. ■

Remark 6: Note that the graph condition "$\mathcal{D}_h$ is strongly $(f+1)$-robust w.r.t. $\mathcal{L}_h$" in Lemma 1 is somewhat idealized, since it requires the identities of the healthy leaders to be known *a priori*. To render the graph condition more practical, we adopt a similar idea of [28], which assumes that a small number of healthy agents cannot be attacked (defined as trusted agents) and forms a robust subgraph among these agents. Likewise, we assume a small group of leaders to be trusted (denoted as $\mathcal{L}_t$). Unlike healthy agents, these trusted leaders adhere to predetermined update rules, but cannot be compromised by adversarial attacks. Subsequently, we define a subgraph $\mathcal{D}_t$ that captures the agents in $\mathcal{L}_t \cup \mathcal{F}_h$ and the connections between these agents. Then, the original graph



(a) The idea of constructing a robust graph with trusted agents in [28]. The constructed graph is an undigraph under the leaderless structure.



(b) The idea of constructing a strongly robust graph with trusted leaders in this paper. The constructed graph is a digraph under the leader-follower structure.

Fig. 1. Illustration of constructing robust graphs with trusted agents.

condition is modified to: "$\mathcal{D}_t$ is strongly $(f+1)$-robust w.r.t. $\mathcal{L}_t$". This modified graph condition can also serve as a sufficient graph condition for achieving leader-follower resilient rendezvous, while it only needs to enhance the security of a small number of nodes. Thus, in practical scenarios, we can use the modified condition to construct strongly robust graphs. Figs. 1(a) and (b) illustrate the ideas of constructing robust graphs with trusted agents in [28] and in this paper, respectively. Nonetheless, the original graph condition is still adopted in this paper due to its significance in theoretical analysis and practical scenarios.

With Lemma 1, we further demonstrate the monotonicity of $\overline{M}[t]$ and $\underline{m}[t]$ as follows.

Lemma 2: Suppose the same preconditions as in Lemma 1. For all time steps $t \geq 0$, the following statements hold:

1) $\overline{M}[t]$ is nonincreasing and $\underline{m}[t]$ is nondecreasing.
2) $x_i[t+1] \in \left[\underline{m}[t], \overline{M}[t]\right], \ \forall i \in \mathcal{H}$.

Proof: Here we only provide the proof for healthy followers, since the healthy leaders are static and their state values always satisfy the above two statements.

We start with the proof of monotonicity for $\overline{M}[t]$ and $\underline{m}[t]$. Consider a healthy follower $i \in \mathcal{F}_h$. By invoking the update rule (6), Assumption 1, and Lemma 1, the state $x_i[t+1]$ is upper bounded by

$$\begin{aligned} x_i[t+1] &\leq \theta_{ii}[t]\overline{M}[t] + \sum_{j \in \mathcal{N}_i^{\text{in}}[t] \setminus \mathcal{R}_i^{\text{in}}[t]} \theta_{ij}[t]\overline{M}[t] \\ &= \overline{M}[t]. \end{aligned} \tag{8}$$

Note that the above relation holds for all $i \in \mathcal{F}_h$, which yields $\overline{M}[t+1] \leq \overline{M}[t]$. Likewise, we have $\underline{m}[t+1] \geq \underline{m}[t]$. These facts indicate the monotonicity of $\overline{M}[t]$ and $\underline{m}[t]$, which completes the proof of the first statement of Lemma 1.

Now, let us turn our attention to the second statement. It follows from (6) that

$$\begin{aligned} x_i[t+1] &\geq \theta_{ii}[t]\underline{m}[t] + \sum_{j \in \mathcal{N}_i^{\text{in}}[t] \setminus \mathcal{R}_i^{\text{in}}[t]} \theta_{ij}[t]\underline{m}[t] \\ &= \underline{m}[t]. \end{aligned} \tag{9}$$

Synthesizing (8) with (9), we obtain $\underline{m}[t] \leq x_i[t+1] \leq \overline{M}[t]$. Note that this relation holds for all $t \geq 0$ and for all $i \in \mathcal{F}_h$. Thus, we eventually reach $x_i[t+1] \in \left[\underline{m}[t], \overline{M}[t]\right], \ \forall i \in \mathcal{H}, \ \forall t \geq 0$. ∎

With these results, it is time to give the condition on the network topology to guarantee leader-follower resilient rendezvous with the LFAT algorithm.

Theorem 1: Consider a leader-follower MAS described by a time-invariant digraph $\mathcal{D} = (\mathcal{V}, \mathcal{E})$. Let Assumption 1 hold. Suppose that the misbehaving set $\mathcal{M}$ satisfies the f-local attack model stated in Definition 7. Then, the necessary and sufficient condition for achieving resilient rendezvous with the LFAT algorithm is that $\mathcal{D}_h$ is strongly $(f+1)$-robust w.r.t. $\mathcal{L}_h$.

Proof (Necessity): We prove the necessity with a contradiction argument. If $\mathcal{D}_h$ is not strongly $(f+1)$-robust w.r.t. $\mathcal{L}_h$, then there exists a healthy follower $i \in \mathcal{F}_h$ who possesses at most f healthy leaders (with the state values Υ) and at most f misbehaving agents as in-neighbors. Suppose that the MAS has already reached rendezvous, i.e., all healthy followers reach rendezvous on Υ. Through the LFAT algorithm, at most $f-1$ largest and smallest in-neighbor values will be filtered out. If the misbehaving agents maintain their state values at $c \in \mathbb{R}, c \neq \Upsilon$, at least one misbehaving agent will be retained, which results in a breach of rendezvous.

(Sufficiency): For any $\mu \in \mathbb{R}$ and $t, \ell \in \mathbb{N}$, let

$$\begin{aligned} \mathcal{X}_{\overline{M}}(t,\ell,\mu) &= \left\{j \in \mathcal{F} : x_j[t+\ell] > \overline{M}[t] - \mu\right\}, \\ \mathcal{X}_{\underline{m}}(t,\ell,\mu) &= \left\{j \in \mathcal{F} : x_j[t+\ell] < \underline{m}[t] + \mu\right\}. \end{aligned} \tag{10}$$

Furthermore, we define

$$\begin{aligned} \mathcal{Y}_{\overline{M}}(t,\ell,\mu) &= \mathcal{X}_{\overline{M}}(t,\ell,\mu) \cap \mathcal{H}, \\ \mathcal{Y}_{\underline{m}}(t,\ell,\mu) &= \mathcal{X}_{\underline{m}}(t,\ell,\mu) \cap \mathcal{H}, \end{aligned} \tag{11}$$

which capture the healthy agents in $\mathcal{X}_{\overline{M}}(t,\ell,\mu)$ and $\mathcal{X}_{\underline{m}}(t,\ell,\mu)$, respectively. Next, we define

$$\mathcal{U}_{\mathcal{Y}}(t,\ell,\mu)=\mathcal{Y}_{\overline{M}}(t,\ell,\mu)\cup\mathcal{Y}_{\underline{m}}(t,\ell,\mu) \tag{12}$$

as the union of the two sets in (11). In addition, let

$$\begin{aligned} D[t]&=\overline{M}[t]-\underline{m}[t],\\ \mu(t,0)&=(\overline{M}[t]-\underline{m}[t])/2,\\ \mu(t,\ell)&=\rho^{\ell}\mu(t,0). \end{aligned} \tag{13}$$

With these notions, we first concentrate on $\mathcal{Y}_{\overline{M}}(t,0,\mu(t,0))$ and $\mathcal{Y}_{\underline{m}}(t,0,\mu(t,0))$. Consider the set $\mathcal{U}_0=\mathcal{F}_h$. Since $\mathcal{D}_h$ is strongly $(f+1)$-robust w.r.t. $\mathcal{L}_h$ and the misbehaving set $\mathcal{M}$ fulfills the f-local attack model, there exists at least one healthy follower $i\in\mathcal{F}_h$ in $\mathcal{Y}_{\overline{M}}(t,0,\mu(t,0))$ or $\mathcal{Y}_{\underline{m}}(t,0,\mu(t,0))$ that has at least $f+1$ healthy in-neighbors outside its respective set and at most f in-neighbors from $\mathcal{M}$.

We first suppose $i\in\mathcal{Y}_{\overline{M}}(t,0,\mu(t,0))$. Considering the implementation of the LFAT algorithm, at least one in-neighbor $j\in\mathcal{N}_i^{\text{in}}[t]\backslash\mathcal{R}_i^{\text{in}}[t]$ satisfies $x_j^i[t]\leq\overline{M}[t]-\mu(t,0)$ and this $x_j^i[t]$ is remained for state update. Note that the state update of agent i is a convex combination of its own state and the state it uses from its in-neighbors, and each coefficient in the combination is lower bounded by ρ due to Assumption 1. From Lemma 2, one knows that the largest state that agent i will use at time step t is $\overline{M}[t]$. By invoking the convexity inequality, placing the largest possible weight $1-\rho$ on $\overline{M}[t]$, and placing the smallest possible weight ρ on $\overline{M}[t]-\mu(t,0)$, one produces

$$\begin{aligned} x_i[t+1]&\leq(1-\rho)\overline{M}[t]+\rho(\overline{M}[t]-\mu(t,0))\\ &=\overline{M}[t]-\rho\mu(t,0)=\overline{M}[t]-\mu(t,1), \end{aligned} \tag{14}$$

which means that agent i will move out from $\mathcal{Y}_{\overline{M}}(t,1,\mu(t,1))$ at time step $t+1$ since $x_i[t+1]\leq\overline{M}[t]-\mu(t,1)$. Note that the relation (14) also holds for healthy followers in $\mathcal{F}_h\backslash\mathcal{X}_{\overline{M}}(t,0,\mu(t,0))$. Thus, one obtains

$$|\mathcal{Y}_{\overline{M}}(t,1,\mu(t,1))|<|\mathcal{Y}_{\overline{M}}(t,0,\mu(t,0))|. \tag{15}$$

If the healthy follower $i\in\mathcal{Y}_{\underline{m}}(t,0,\mu(t,0))$, the implementation of the LFAT algorithm ensures that at least one of agent i's in-neighbor $j\in\mathcal{N}_i^{\text{in}}[t]\backslash\mathcal{R}_i^{\text{in}}[t]$ satisfies $x_j^i[t]\geq\underline{m}[t]+\mu(t,0)$ and this $x_j^i[t]$ is remained for state update. From Lemma 2, one knows that the smallest state that agent i will use at time step t is $\underline{m}[t]$. By invoking the convexity inequality, placing the largest possible weight $1-\rho$ on $\underline{m}[t]$, and placing the smallest possible weight ρ on $\underline{m}[t]+\mu(t,0)$, one has

$$\begin{aligned} x_i[t+1]&\geq(1-\rho)\underline{m}[t]+\rho(\underline{m}[t]+\mu(t,0))\\ &=\underline{m}[t]+\rho\mu(t,0)=\underline{m}[t]+\mu(t,1), \end{aligned} \tag{16}$$

which means that agent i will move out from $\mathcal{Y}_{\underline{m}}(t,1,\mu(t,1))$ at time step $t+1$ since $x_i[t+1]\geq\underline{m}[t]+\mu(t,1)$. Note that the relation (16) also applies to healthy followers in $\mathcal{F}_h\backslash\mathcal{X}_{\underline{m}}(t,0,\mu(t,0))$. Thus, one has

$$|\mathcal{Y}_{\underline{m}}(t,1,\mu(t,1))|<|\mathcal{Y}_{\underline{m}}(t,0,\mu(t,0))|. \tag{17}$$

Synthesizing (15) with (17), one concludes that no matter whether agent i belongs to $\mathcal{Y}_{\overline{M}}(t,0,\mu(t,0))$ or $\mathcal{Y}_{\underline{m}}(t,0,\mu(t,0))$, it holds

$$\begin{aligned} &|\mathcal{Y}_{\overline{M}}(t,1,\mu(t,1))|+|\mathcal{Y}_{\underline{m}}(t,1,\mu(t,1))|\\ <&|\mathcal{Y}_{\overline{M}}(t,0,\mu(t,0))|+|\mathcal{Y}_{\underline{m}}(t,0,\mu(t,0))|. \end{aligned} \tag{18}$$

According to the definition in (12), the relation (18) can be rewritten as

$$|\mathcal{U}_{\mathcal{Y}}(t,1,\mu(t,1))|<|\mathcal{U}_{\mathcal{Y}}(t,0,\mu(t,0))|, \tag{19}$$

since $\mathcal{Y}_{\overline{M}}(t,0,\mu(t,0))$ and $\mathcal{Y}_{\underline{m}}(t,0,\mu(t,0))$ are disjoint, and $\mathcal{Y}_{\overline{M}}(t,1,\mu(t,1))$ and $\mathcal{Y}_{\underline{m}}(t,1,\mu(t,1))$ are disjoint.

Next, we consider the set $\mathcal{U}_1=\mathcal{U}_{\mathcal{Y}}(t,1,\mu(t,1))$. If $\mathcal{U}_1$ is not empty, we adopt a similar analysis as above and obtain that there exists at least one healthy follower $i_1\in\mathcal{F}_h$ in $\mathcal{Y}_{\overline{M}}(t,1,\mu(t,1))$ or $\mathcal{Y}_{\underline{m}}(t,1,\mu(t,1))$ that has at least $f+1$ healthy in-neighbors outside its respective set and at most f in-neighbors from $\mathcal{M}$. We first suppose $i_1\in\mathcal{Y}_{\overline{M}}(t,1,\mu(t,1))$. Then, the state $x_{i_1}[t+2]$ follows

$$\begin{aligned} x_{i_1}[t+2]&\leq(1-\rho)\overline{M}[t+1]+\rho(\overline{M}[t]-\mu(t,1))\\ &\leq\overline{M}[t]-\rho^2\mu(t,0)=\overline{M}[t]-\mu(t,2). \end{aligned} \tag{20}$$

Note that the above relation also holds for healthy followers in $\mathcal{F}_h\backslash\mathcal{X}_{\overline{M}}(t,1,\mu(t,1))$. If $i_1\in\mathcal{Y}_{\underline{m}}(t,1,\mu(t,1))$, one has

$$\begin{aligned} x_{i_1}[t+2]&\geq(1-\rho)\underline{m}[t+1]+\rho(\underline{m}[t]+\mu(t,1))\\ &\geq\underline{m}[t]+\rho^2\mu(t,0)=\underline{m}[t]+\mu(t,2). \end{aligned} \tag{21}$$

The relation (21) also applies to healthy followers in $\mathcal{F}_h\backslash\mathcal{X}_{\underline{m}}(t,1,\mu(t,1))$. Consequently, it yields

$$|\mathcal{U}_{\mathcal{Y}}(t,2,\mu(t,2))|<|\mathcal{U}_{\mathcal{Y}}(t,1,\mu(t,1))|. \tag{22}$$

By recursion, we know that for any $\ell\in\mathbb{N}$, it holds

$$|\mathcal{U}_{\mathcal{Y}}(t,\ell,\mu(t,\ell))|<|\mathcal{U}_{\mathcal{Y}}(t,\ell+1,\mu(t,\ell+1))|. \tag{23}$$

Since $|\mathcal{U}_{\mathcal{Y}}(t,0,\mu(t,0))|\leq n$, either $\mathcal{Y}_{\overline{M}}(t,n,\mu(t,n))$ or $\mathcal{Y}_{\underline{m}}(t,n,\mu(t,n))$ (or both) will be empty. Regarding the former case, it holds $\overline{M}[t+n]\leq\overline{M}[t]-\mu(t,n)$. By invoking the first statement of Lemma 2, it yields $\underline{m}[t+n]\geq\underline{m}[t]$. Synthesizing these two results yields

$$\begin{aligned} D[t+n]&=\overline{M}[t+n]-\underline{m}[t+n]\\ &\leq\overline{M}[t]-\mu(t,n)-\underline{m}[t]\\ &=(1-\frac{\rho^n}{2})D[t]. \end{aligned} \tag{24}$$

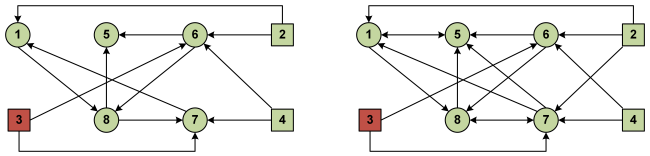
Note that the relation (24) holds for any $t\in\mathbb{N}$. Thus, we further derive

$$D[t+kn]\leq(1-\frac{\rho^n}{2})^kD[t],\ \forall k\in\mathbb{N}. \tag{25}$$

Since $1-\frac{\rho^n}{2}\in(0,1)$, we eventually arrive at $\lim_{k\to\infty}D[t+kn]=0$, which implies $\lim_{t\to\infty}D[t]=0$. This fact means that rendezvous among the healthy agents is achieved, i.e., $\lim_{t\to\infty}\overline{M}[t]=\lim_{t\to\infty}\underline{m}[t]$.

It remains to prove that all healthy followers will reach rendezvous on Υ. By contradiction, we assume $\lim_{t\to\infty}\overline{M}[t]=\lim_{t\to\infty}\underline{m}[t]\neq\Upsilon$. By the condition of strong $(f+1)$-robustness, there exists at least one healthy follower $i\in\mathcal{F}_h$ who possesses at least $f+1$ healthy leaders as in-neighbors. Through the LFAT algorithm, at least one leader with the state value Υ is retained and the rendezvous state is violated. Thus, we eventually arrive at $\lim_{t\to\infty}\overline{M}[t]=\lim_{t\to\infty}\underline{m}[t]=\Upsilon$, which implies that resilient rendezvous is achieved. ■

For the convenience of subsequent comparison, we present the convergence result using the W-MSR as follows.



(a) LFAT. (b) W-MSR.

Fig. 2. Digraphs fulfilling sufficient conditions for achieving resilient rendezvous with different resilient algorithms ($f = 1$).

Theorem 2 ([13]): Consider a leader-follower MAS described by the digraph $\mathcal{D} = (\mathcal{V}, \mathcal{E})$. Let Assumption 1 hold. Suppose that the misbehaving set $\mathcal{M}$ satisfies the f-local attack model stated in Definition 7. Then, a sufficient condition to achieve a resilient rendezvous with the W-MSR algorithm is that $\mathcal{D}$ is strongly $(2f+1)$-robust w.r.t. $\mathcal{L}$.

Remark 7: Figs. 2(a) and (b) show the two sufficient graph conditions to achieve leader-follower resilient rendezvous with the LFAT and W-MSR algorithms (denoted as Condition 1 and Condition 2, respectively). Their main differences are summarized as follows:

1) Prior information. Both of them require the identities of leaders and followers as prior information, while Condition 1 needs to know the identities of all healthy leaders in advance;
2) Graph construction emphasis. The construction of Condition 1 focuses on the robustness of local healthy network, while the construction of Condition 2 concentrates on the robustness of overall network;
3) Network redundancy. Compared with Condition 2, Condition 1 can reduce at most 50% of the network redundancy in the absence of adversarial attacks by relaxing the graph condition from strong $(2f+1)$-robustness to strong $(f+1)$-robustness. When subject to the same attack model, the network redundancy of Condition 1 is also lower than that of Condition 2.

In addition, Condition 2 is a stricter graph condition than Condition 1, i.e., if the overall network $\mathcal{D}$ is $(2f+1)$-robust w.r.t. $\mathcal{L}$, then the healthy network $\mathcal{D}_h$ is strongly $(f+1)$-robust w.r.t. $\mathcal{L}_h$, but not vice versa.

Note that the above conclusions can be extended to time-varying networks with the help of Definition 4.

Corollary 1: Consider a leader-follower MAS modeled by a time-varying digraph $\mathcal{D}[t] = (\mathcal{V}, \mathcal{E}[t])$. Let Assumption 1 hold. Suppose that the misbehaving set $\mathcal{M}$ satisfies the f-local attack model stated in Definition 7. Then, a sufficient condition for achieving resilient rendezvous with the LFAT algorithm is that $\mathcal{D}_h[t]$ is jointly and strongly $(f+1)$-robust w.r.t. $\mathcal{L}_h$.

Proof: Similar to the proof of Theorem 1, one constructs two nonempty and disjoint sets $\mathcal{Y}_{\overline{M}}(t, t_h - t, \mu(t, t_h - t))$ and $\mathcal{Y}_{\underline{m}}(t, t_h - t, \mu(t, t_h - t))$. As the underlying network $\mathcal{D}_h[t]$ of healthy agents is jointly and strongly $(f+1)$ robust w.r.t. $\mathcal{L}_h$, there exists a time step t_j in each time interval $[t_h, t_{h+1})$ such that there exists a healthy follower $i \in \mathcal{F}_h$ in $\mathcal{Y}_{\overline{M}}(t, t_h - t, \mu(t, t_h - t))$ or $\mathcal{Y}_{\underline{m}}(t, t_h - t, \mu(t, t_h - t))$ that has at least $f+1$ healthy neighbors outside of its respective set. Subsequently, one derives

$$|\mathcal{U}_{\mathcal{Y}}(t, t_{h+1} - t, \mu(t, t_{h+1} - t))| < |\mathcal{U}_{\mathcal{Y}}(t, t_h - t, \mu(t, t_h - t))| \tag{26}$$

if both $\mathcal{Y}_{\overline{M}}(t, t_h - t, \mu(t, t_h - t))$ and $\mathcal{Y}_{\underline{m}}(t, t_h - t, \mu(t, t_h - t))$ are nonempty. Therefore, at least one of $\mathcal{Y}_{\overline{M}}(t, nT, \mu(t, nT))$ and $\mathcal{Y}_{\underline{m}}(t, nT, \mu(t, nT))$ is empty due to $|\mathcal{Y}_{\overline{M}}(t, 0, \mu(t, 0))| + \left|\mathcal{Y}_{\underline{m}}(t, 0, \mu(t, 0))\right| \leq n$, where T is the maximum length of time intervals $[t_h, t_{h+1})$. For both of these two cases, one shows

$$\begin{aligned} D[t + nT] &\leq D[t] - \mu(t, nT) \\ &= (1 - \frac{\rho^{nT}}{2}) D[t], \end{aligned} \tag{27}$$

which results in $\lim_{k\to\infty} D[t + knT] = 0$. Thus, we eventually reach $\lim_{t\to\infty} D[t] = 0$.

It remains to prove that the state values of all healthy followers converge to Υ. By adopting a similar analysis to Theorem 1, we eventually arrive at $\lim_{t\to\infty} \overline{M}[t] = \lim_{t\to\infty} \underline{m}[t] = \Upsilon$. This completes the proof of resilient rendezvous for leader-follower MASs in time-varying networks. ■

IV. Numerical Example

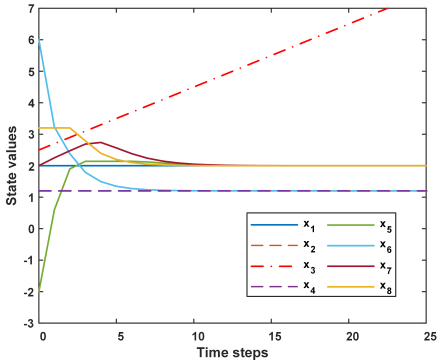
In this section, several numerical examples are given to validate that strong $(f+1)$-robustness is necessary and sufficient to achieve resilient rendezvous for leader-follower MASs. The superiority and scalability of the LFAT algorithm are also validated, respectively.

A. Comparison Between the LFAT and W-MSR Algorithms

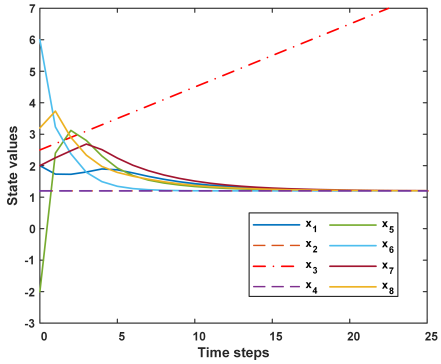
1) Under Digraph Fig. 2(a): We consider a digraph $\mathcal{D} = (\mathcal{V}, \mathcal{E})$ with $\mathcal{L} = \{2, 3, 4\}$ and $\mathcal{F} = \{1, 5, 6, 7, 8\}$; see Fig. 2(a). Note that $\mathcal{D}_h$ is strongly 2-robust w.r.t. $\mathcal{L}_h$, which satisfies the sufficient condition for achieving leader-follower resilient rendezvous with the LFAT algorithm when the MAS is subject to the 1-local attack model, but not for the W-MSR algorithm. In addition, we assume that Leader 3 is misbehaving, whose motion is expressed as $x_3[t] = 2.5 + 0.2 \times t$. It can be verified that the MAS satisfies the 1-local attack model. Let the initial states of all agents be $[x_1[0], \cdots, x_8[0]]^{\mathrm{T}} = [2, 1.2, 2.5, 1.2, -2, 6, 2, 3.2]^{\mathrm{T}}$, which implies $\Upsilon = 1.2$.

To start with, we apply the W-MSR algorithm to the leader-follower MAS. The convergence result is depicted in Fig. 3(a). Notice that only the state value of Agent 6 converges to the right value $\Upsilon = 1.2$, while other healthy followers fail to achieve leader-follower resilient rendezvous. Note that the state value of Agent 1 remains unchanged. This is because the W-MSR algorithm removes all available in-neighbors of Agent 1 used for state updates.

As a comparison, the convergence result with the LFAT algorithm is depicted in Fig. 3(b), from which we observe that all healthy followers reach rendezvous on $\Upsilon = 1.2$ despite the influence of misbehaving Agent 3. The implementation of the LFAT algorithm ensures that each healthy follower has a certain number of in-neighbors to update its own state. These results also imply that, compared to the W-MSR algorithm, the proposed LFAT algorithm enables the MAS to achieve leader-follower resilient rendezvous with relaxed graph conditions.



(a) The state values of healthy followers fail to achieve leader-follower resilient rendezvous.



(b) The state values of all healthy followers achieve leader-follower resilient rendezvous.

Fig. 3. Convergence results of the MAS under Fig. 2(a) with (a) the W-MSR algorithm and (b) the LFAT algorithm. The solid lines indicate the healthy followers, the dotted lines indicate the healthy leaders, and the dashed dot line indicates the misbehaving agent.

2) Under Digraph Fig. 2(b): In this case, the digraph $\mathcal{D}$ is strongly 3-robust w.r.t. $\mathcal{L}$, which implies that each follower has at least 3 in-neighbors. By invoking (5), it holds $\varepsilon = f$ and the LFAT algorithm becomes the W-MSR algorithm. Consequently, there is no need to compare the two algorithms in this situation, since they will exhibit the identical performance.

3) Under Respective Sufficient Graph Conditions: To further compare the performance of the LFAT and W-MSR algorithms using different metrics, we let the two algorithms operate in their respective sufficient graph conditions, i.e., the LFAT algorithm operates in Fig. 2(a) and the the W-MSR algorithm operates in Fig. 2(b). Other settings are the same as that in *Case 1)*.

The convergence result under Fig. 2(a) with the LFAT algorithm has been illustrated in Fig. 3(b), while the convergence result under Fig. 2(b) with the W-MSR algorithm is depicted in Fig. 4. It can be observed that resilient rendezvous is achieved in both situations. However, by calculating the number of directed edges of Fig. 2(a) and Fig. 2(b), respectively, we find that the network redundancy using the LFAT algorithm is lower than that of W-MSR algorithm. This is because the propose method relaxes the graph robustness requirement. To show the convergence rate of the leader-follower MAS using the LFAT and W-MSR algorithms more intuitively, we define

$$e[t] = \left| \max_{i \in \mathcal{F}_h} \{x_i[t]\} - \Upsilon \right| \tag{28}$$

as the convergence error. According to Definition 8, the leader-follower MAS will achieve resilient rendezvous if $e[t]$ asymptotically approaches zero. Then, the convergence errors with the LFAT and W-MSR algorithms under their respective sufficient graph conditions are shown in Fig. 5, from which we observe that the LFAT algorithm demonstrates a faster convergence rate. This is because the LFAT algorithm retains more available information. To quantify the result, we define the time step when $e[t] \leq 0.001$ is first reached as the convergence time. Then, the above quantitative results, including the number of directed edges and the convergence time of the two algorithms are listed in TABLE I, from which we observe that the proposed LFAT algorithm outperforms the existing

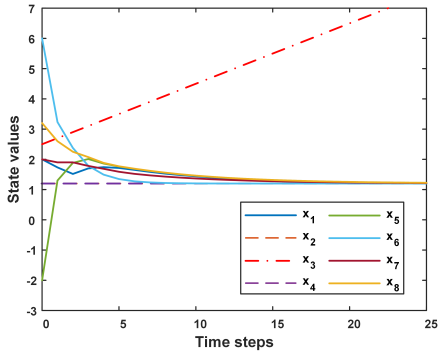


Fig. 4. Convergence result of the MAS under Fig. 2(b) with the W-MSR algorithm.

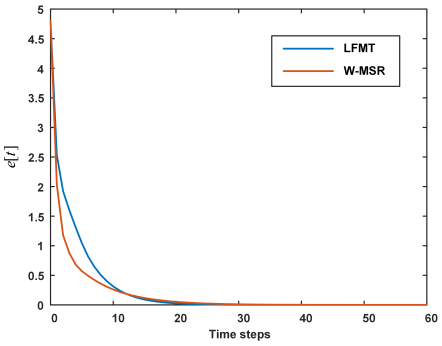


Fig. 5. Convergence errors with the LFAT and W-MSR algorithms under their respective sufficient graph conditions.

W-MSR algorithm in terms of the network redundancy and convergence rate, since it shares fewer directed edges and a shorter convergence time.

B. Resilient Rendezvous in Large-Scale Networks

To validate the scalability of the LFAT algorithm, we extend the obtained results to a larger-scale network. To this end, we consider a strongly 2-robust digraph with sixteen nodes and described by Fig. 6, where $\mathcal{L} = \{4, 5, \ldots, 8\}$ and the remaining agents are followers. Furthermore, we let Leader 8

TABLE I
PERFORMANCE COMPARISON BETWEEN THE W-MSR ALGORITHM AND THE LFAT ALGORITHM UNDER THEIR RESPECTIVE SUFFICIENT GRAPH CONDITIONS.

Algorithm	Sufficient Graph Condition	Directed Edges	Convergence Time
W-MSR	$\mathcal{D}$ is strongly $(2f+1)$-robust w.r.t. $\mathcal{L}$	17	Time step 46
LFAT	$\mathcal{D}_h$ is strongly $(f+1)$-robust w.r.t. $\mathcal{L}_h$	12	Time step 34

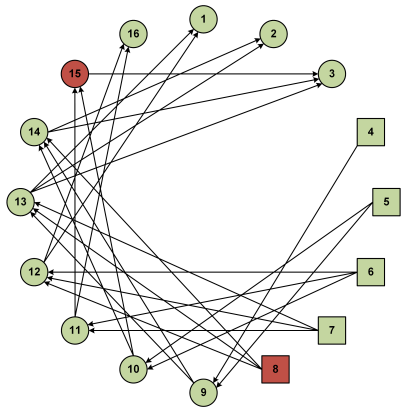


Fig. 6. A strongly 2-robust digraph with sixteen agents.

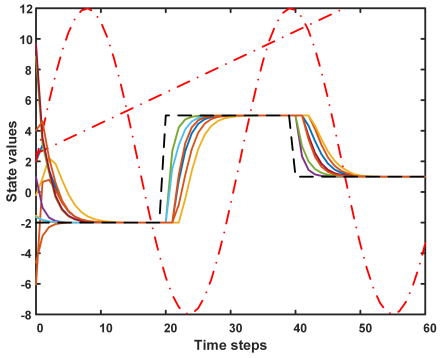


Fig. 7. State values of agents in a larger-scale MAS. The trajectories of healthy followers (solid line) converge to the reference value (black dashed line) propagated by healthy leaders, despite two misbehaving agents (red dashed dot line) exist in the network.

and Follower 15 be misbehaving agents, whose trajectories are described as $x_8[t] = 2.5+0.2\cdot t$ and $x_{15}[t] = 10\cdot\sin(t/5)+2$, respectively. Although there exist two misbehaving agents, the MAS still satisfies the 1-local attack model according to Definition 7. The initial states of all healthy followers are random values in the interval $[-10, 10]$, while the state value Υ of the leaders is chosen as

$$\Upsilon = \begin{cases} -2, & \text{if } t \in [0, 20), \\ 5, & \text{if } t \in [20, 40), \\ 1, & \text{if } t \in [40, 60], \end{cases} \tag{29}$$

The state values of all agents using the LFAT algorithm are shown in Fig. 7, from which we observe that the state values of healthy followers (solid line) successfully converge to the reference value (black dashed line), despite the influence of two misbehaving agents (red dashed dot line) in the network. Therefore, resilient rendezvous is guaranteed. This simulation also demonstrates that the proposed LFAT algorithm is scalable for large-scale networks.

C. Resilient Rendezvous in Time-Varying Networks

In this subsection, resilient rendezvous is further guaranteed in time-varying networks to relax the graph robustness requirement at each time step. To this end, we consider a jointly and strongly 2-robust digraph $\mathcal{D}[t]$ with sixteen agents, where $\mathcal{L} = \{4, 5, \ldots, 8\}$ and the remaining agents are followers. The state value Υ of the leaders is chosen as $\Upsilon = -2$. The network topology is shown in Fig. 8 and switches as follows:

$$\mathcal{D}[t] = \begin{cases} \mathcal{D}_1, & t = 3m, \\ \mathcal{D}_2, & t = 3m+1, \\ \mathcal{D}_3, & t = 3m+2. \end{cases} \quad m \in \mathbb{N}. \tag{30}$$

Other simulation settings are the same as that in Section IV-B.

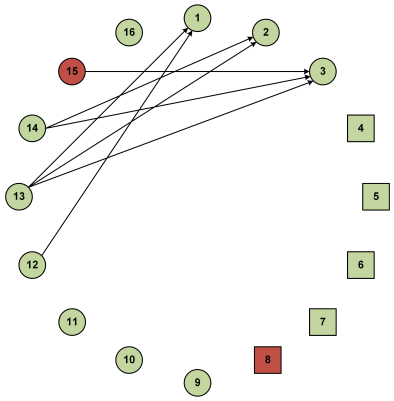
Fig. 9 illustrates the state values of sixteen agents with the proposed LFAT algorithm in time-varying networks, respectively. Although the misbehaving agents 8 and 15 deviate from the normal state update, all healthy followers successfully reach rendezvous on the reference value propagated by healthy leaders. These results validate Corollary 1. Furthermore, it is worth mentioning that, though none of the three digraphs in Fig. 8 is a strongly robust graph, resilient rendezvous is still ensured for leader-follower MASs under the 1-local attack model. This fact means that the conventional requirement on network topology is relaxed by introducing the jointly and strongly robust graph. These results also exhibit that the proposed LFAT algorithm is scalable for time-varying networks.

D. Application of the LFAT Algorithm to Microgrid Systems

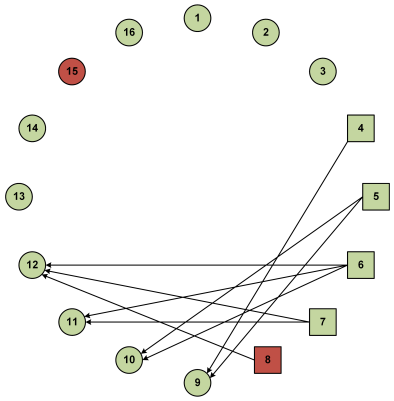
In this subsection, we apply the proposed LFAT algorithm to a single-phase microgrid system and endeavor to tackle a frequency synchronization problem when the system is subject to adversarial attacks. Fig. 10 illustrates the simulated microgrid [38], which consists of twenty voltage source inverters serving as grid-forming generators and twenty points of common coupling (PCCs) to connect the generator to adjacent power grids. The corresponding communication network of the microgrid is depicted in Fig. 11, which implies that the system operates in a leader-follower mode with Generators 1 and 2 being the leaders and the remaining generators being followers. Furthermore, we consider that the microgrid is suffering from an external adversarial attack, where all generators operate normally and two attackers try to compromise Generators 6, 10, 11, and 15. Specifically, the attackers disguise themselves as in-neighbors of these generators and inject false data into them at each time step.

Regarding the frequency synchronization problem [39], the microgrid is to be synchronized at points of interconnection near Generators 1 and 2. Then, the frequency reference correction $\Delta\omega_l,\ l = 1, 2$ of Generators 1 and 2 is expressed as the difference between the measured frequency $\omega_{\text{PCC}l}$ of the microgrid and the nominal reference ω^*:

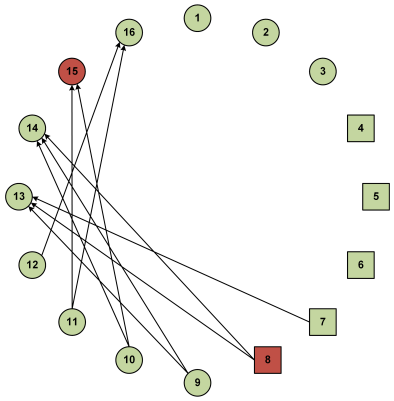
$$\Delta\omega_l = \omega_{\text{PCC}l} - \omega^*, \quad l = 1, 2, \tag{31}$$



(a) Digraph $\mathcal{D}_1$.



(b) Digraph $\mathcal{D}_2$.



(c) Digraph $\mathcal{D}_3$.

Fig. 8. A jointly and strongly 2-robust digraph with sixteen agents.

while the frequency reference corrections of the remaining generators adhere to the resilient strategy modified from [39]:

$$\Delta\omega_i[t+1] = \left(1 - \frac{1}{k_s} \sum_{j \in \mathcal{N}_i^{\text{in}}[t] \setminus \mathcal{R}_i^{\text{in}}[t]} \theta_{ij}[t]\right) \Delta\omega_i[t] + \frac{1}{k_s} \sum_{j \in \mathcal{N}_i^{\text{in}}[t] \setminus \mathcal{R}_i^{\text{in}}[t]} \theta_{ij}[t] \Delta\omega_j[t], \tag{32}$$

where k_s is the distributed synchronization gain. From (31) and (32), it can be observed that the frequency reference corrections follow a leader-follower mode with Generators 1 and 2 being leaders. By implementing the proposed LFAT algorithm, the frequency references of all other generators will converge to the frequency of the microgrid system at Generators 1 and 2 despite the influence of two external attackers.

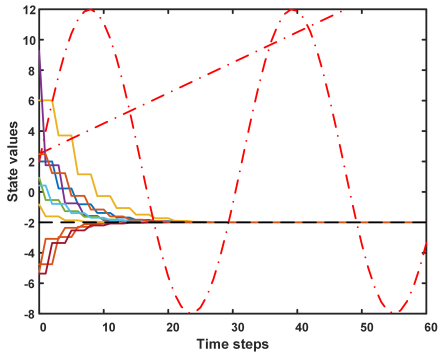


Fig. 9. State values of agents in a time-varying digraph. The trajectories of healthy followers (solid line) converge to the reference value (black dashed line) propagated by healthy leaders, despite two misbehaving agents (red dashed dot line) exist in the network.

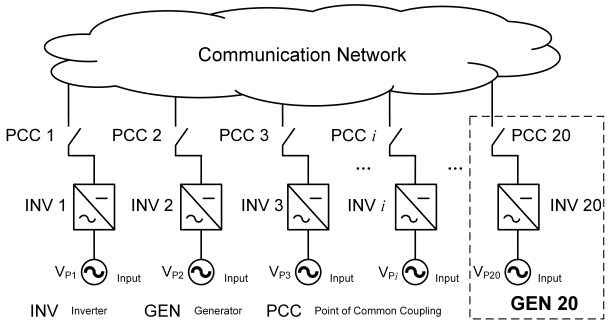


Fig. 10. Illustration of the simulated microgrid.

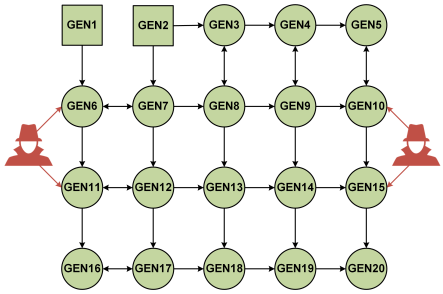


Fig. 11. Communication network of the microgrid.

In the simulation, the frequency reference corrections of all generators are initialized with zero and the gain k_s is set as $k_s = 0.8$. At the 20^{th} time step, the information states $\Delta\omega_1$ and $\Delta\omega_2$ reduce from 0 to -2π rad/s, and the LFAT algorithm is implemented. At the 80^{th} time step, the information states $\Delta\omega_1$ and $\Delta\omega_2$ increase from 0 to 2π rad/s, and the W-MSR algorithm is implemented. The convergence result is shown in Fig. 12. During the 0-80^{th} time steps, the proposed LFAT algorithm effectively keeps $\Delta\omega_i$ rendezvous, thereby achieving distributed frequency synchronization. However, after the 80^{th} time step, not all frequency reference corrections of Generators 3 to 20 are able to follow the changes of Generators 1

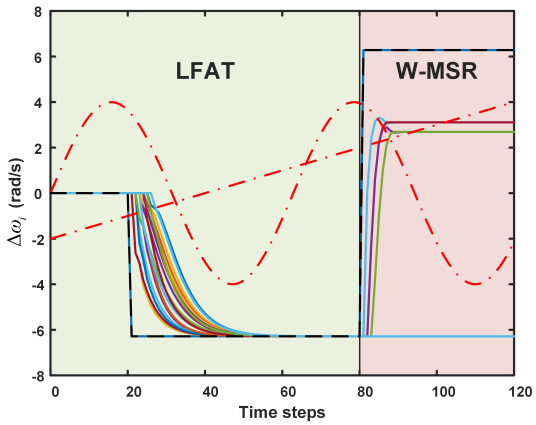


Fig. 12. Convergence result of the microgrid. The solid lines indicate Generators 3-20, the black dotted lines indicate Generators 1 and 2, and the red dashed dot lines indicate the two external attackers.

and 2. Some remain unchanged at -2π rad/s while others cannot converge to 2π rad/s. This simulation demonstrates that frequency synchronization is not guaranteed with the W-MSR algorithm.

V. Conclusion

In this study, we have investigated the leader-follower resilient rendezvous problem for MASs subject to adversarial attacks. Compared with the existing studies, the advantages of the proposed method are as follows.

1) A novel attack-tolerant algorithm. This paper proposes a leader-follower attack-tolerant (LFAT) algorithm, which enables the healthy followers to achieve resilient rendezvous with reduced network redundancy.
2) More complete theoretical conditions. The theoretical part derives the necessary and sufficient condition for achieving resilient rendezvous, which bridges the gap on graph condition analysis.
3) Application to the microgrid system. We apply the proposed LFAT algorithm to a large-scale microgrid system and address a frequency synchronization problem when the system is subject to adversarial attacks. The result shows that the LFAT algorithm effectively keeps the frequency reference corrections rendezvous, thereby achieving distributed frequency synchronization.

Although the proposed LFAT algorithm has been demonstrated to efficiently defend against adversarial attacks with reduced network redundancy, it still has some limitations. For one thing, it requires each healthy follower to communicate with its in-neighbors at each time step to obtain their states. This requirement means that frequent information interactions between agents will occur, which may cause heavy communication overheads. For another, the proposed method assumes that the attack model is static, while the adversarial attack may be dynamically changing. In the future, we will incorporate the event-triggered mechanism to reduce the communication burden and consider more realistic and dynamic attack models. e.g., the susceptible-infected-recovered (SIR) attack model [40].

References

[1] J. Chen, N. Du, Z. Liao, Y. Cao, H. Meng, Y. Han, Y. Zheng, and Y. Tan, "Energy storage battery parameters identification algorithms of a solar powered communication/remote-sensing uav," *IFAC-PapersOnLine*, vol. 52, no. 24, pp. 41–46, 2019.

[2] N. Zhao, P. Shi, W. Xing, and R. K. Agarwal, "Resilient event-triggered control for networked cascade control systems under denial-of-service attacks and actuator saturation," *IEEE Syst. J.*, vol. 16, no. 1, pp. 1114–1122, 2021.

[3] Y. Zhang, S. Oğuz, S. Wang, E. Garone, X. Wang, M. Dorigo, and M. K. Heinrich, "Self-reconfigurable hierarchical frameworks for formation control of robot swarms," *IEEE Trans. Cybern.*, pp. 1–14, 2023.

[4] Z. Ahmed, N. Ali, and W. Zhang, "H_2 resilient consensus control of multiagent systems under deception attack," *IEEE Trans. Circuits Syst. II-Express Briefs*, vol. 70, no. 7, pp. 2530–2534, 2023.

[5] R. Chen, S. Wang, C. Zhang, H. Dui, Y. Zhang, Z. Yadong, and L. Yang, "Component uncertainty importance measure in complex multi-state system considering epistemic uncertainties," *Chin. J. Aeronaut.*, 2024, doi: 10.1016/j.cja.2024.05.024.

[6] Y. Wang, X. Liu, Y. Kang, and S. S. Ge, "Anomaly resilient relative pose estimation for multiple nonholonomic mobile robot systems," *IEEE Syst. J.*, vol. 16, no. 1, pp. 659–670, 2020.

[7] Z. Liao, S. Wang, J. Shi, M. Li, Y. Zhang, and Z. Sun, "Resilient distributed optimization for cyber–physical systems under adversarial environments: An event-based method," *ISA Trans.*, vol. 149, pp. 1–15, 2024.

[8] N. Zaeri and R. R. Qasim, "Intelligent wireless sensor network for gas classification using machine learning," *IEEE Syst. J.*, 2023.

[9] Y. Zhang, S. Li, S. Wang, X. Wang, and H. Duan, "Distributed bearing-based formation maneuver control of fixed-wing uavs by finite-time orientation estimation," *Aerosp. Sci. Technol.*, vol. 136, pp. 1–13, 2023.

[10] Y. Zhang, S. Wang, X. Wang, and X. Tian, "Bearing-based formation control for multiple underactuated autonomous surface vehicles with flexible size scaling," *Ocean Eng.*, vol. 267, p. 113242, 2023.

[11] Z. Liao, J. Shi, Y. Zhang, S. Wang, and Z. Sun, "A survey of resilient coordination for cyber-physical systems against malicious attacks," *arXiv preprint arXiv:2402.10505*, 2024.

[12] H. J. LeBlanc, H. Zhang, X. Koutsoukos, and S. Sundaram, "Resilient asymptotic consensus in robust networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 4, pp. 766–781, 2013.

[13] J. Usevitch and D. Panagou, "Resilient leader-follower consensus to arbitrary reference values," in *Proc. Amer. Control Conf. (ACC)*. IEEE, 2018, pp. 1292–1298.

[14] H. Rezaee, T. Parisini, and M. M. Polycarpou, "Resiliency in dynamic leader–follower multiagent systems," *Automatica*, vol. 125, p. 109384, 2021.

[15] T. Chen, J. Shan, H. Wen, and S. Xu, "Review of attitude consensus of multiple spacecraft," *Astrodynamics*, vol. 6, no. 4, pp. 329–356, 2022.

[16] M. S. Sadabadi, M. W. S. Atman, A. Aynala, and A. Gusrialdi, "Resilient design of leader-follower consensus against cyber-attacks," *IEEE Trans. Control Netw. Syst.*, vol. 11, no. 2, pp. 1080–1092, 2023.

[17] Z. Li, Z. Duan, and L. Huang, "Leader-follower consensus of multi-agent systems," in *Proc. Amer. Control Conf. (ACC)*. IEEE, 2009, pp. 3256–3261.

[18] W. Ren and Y. Cao, *Distributed Coordination of Multi-Agent Networks: Emergent Problems, Models, and Issues*. Springer, 2011.

[19] S. M. Dibaji, H. Ishii, and R. Tempo, "Resilient randomized quantized consensus," *IEEE Trans. Autom. Control*, vol. 63, no. 8, pp. 2508–2522, 2017.

[20] D. Saldana, A. Prorok, S. Sundaram, M. F. Campos, and V. Kumar, "Resilient consensus for time-varying networks of dynamic agents," in *Proc. Amer. Control Conf. (ACC)*, 2017, pp. 252–258.

[21] J. Yan, C. Wen, X.-K. Liu, and L. Xing, "Resilient impulsive control for second-order consensus under malicious nodes," *IEEE Trans. Circuits Syst. II-Express Briefs*, vol. 68, no. 6, pp. 1962–1966, 2020.

[22] J. Wu, Y. Zhu, Y. Zheng, and H. Wang, "Resilient bipartite consensus of second-order multiagent systems with event-triggered communication," *IEEE Syst. J.*, vol. 17, no. 1, pp. 146–153, 2021.

[23] Y. Bai and J. Wang, "Resilient consensus of continuous-time linear networked systems," *IEEE Trans. Circuits Syst. II-Express Briefs*, vol. 69, no. 8, pp. 3500–3504, 2022.

[24] Z. Liao, J. Shi, S. Wang, Y. Zhang, and Z. Sun, "Resilient consensus through dynamic event-triggered mechanism," *IEEE Trans. Circuits Syst. II-Express Briefs*, vol. 71, no. 7, pp. 3463–3467, 2024.

[25] J. Usevitch and D. Panagou, "Resilient leader-follower consensus with time-varying leaders in discrete-time systems," in *Proc. IEEE Conf. Decis. Control (CDC)*. IEEE, 2019, pp. 5432–5437.

[26] J. Usevitch and D. Panagou, "Resilient leader-follower consensus to arbitrary reference values in time-varying graphs," *IEEE Trans. Autom. Control*, vol. 65, no. 4, pp. 1755–1762, 2019.

[27] J. Yan and C. Wen, "Resilient containment control in adversarial environment," *IEEE Trans. Control Netw. Syst.*, vol. 7, no. 4, pp. 1951–1959, 2020.

[28] W. Abbas, Y. Vorobeychik, and X. Koutsoukos, "Resilient consensus protocol in the presence of trusted nodes," in *Proc. of the 2014 7th International Symposium on Resilient Control Systems (ISRCS)*. IEEE, 2014, pp. 1–7.

[29] W. Abbas, A. Laszka, and X. Koutsoukos, "Improving network connectivity and robustness using trusted nodes with application to resilient consensus," *IEEE Trans. Control Netw. Syst.*, vol. 5, no. 4, pp. 2036–2048, 2017.

[30] Z. Liao, S. Wang, J. Shi, S. Haesaert, Y. Zhang, and Z. Sun, "Resilient containment under time-varying networks with relaxed graph robustness," *IEEE Trans. Netw. Sci. Eng.*, vol. 11, no. 5, pp. 4093–4105, 2024.

[31] W. Fu, J. Qin, Y. Shi, W. X. Zheng, and Y. Kang, "Resilient consensus of discrete-time complex cyber-physical networks under deception attacks," *IEEE Trans. Ind. Inform.*, vol. 16, no. 7, pp. 4868–4877, 2019.

[32] Y. Zhai, Z.-W. Liu, Z.-H. Guan, and G. Wen, "Resilient consensus of multi-agent systems with switching topologies: A trusted-region-based sliding-window weighted approach," *IEEE Trans. Circuits Syst. II-Express Briefs*, vol. 68, no. 7, pp. 2448–2452, 2021.

[33] A. Gusrialdi, "Resilient and privacy-preserving leader-follower consensus in presence of cyber-attacks," *IEEE Control Syst. Lett.*, vol. 7, pp. 3211–3216, 2023.

[34] A. Mitra and S. Sundaram, "Secure distributed observers for a class of linear time invariant systems in the presence of byzantine adversaries," in *Proc. IEEE Conf. Decis. Control (CDC)*. IEEE, 2016, pp. 2709–2714.

[35] G. Wen, Y. Lv, W. X. Zheng, J. Zhou, and J. Fu, "Joint robustness of time-varying networks and its applications to resilient consensus," *IEEE Trans. Autom. Control*, 2023.

[36] W. Ren, R. W. Beard, and E. M. Atkins, "Information consensus in multivehicle cooperative control," *IEEE Control Syst. Mag.*, vol. 27, no. 2, pp. 71–82, 2007.

[37] W. Zhang, T. Qian, X. Chen, K. Huang, W. Tang, and Q. Wu, "Resilient economic control for distributed microgrids under false data injection attacks," *IEEE Trans. Smart Grid*, vol. 12, no. 5, pp. 4435–4446, 2021.

[38] J. W. Simpson-Porco, Q. Shafiee, F. Dörfler, J. C. Vasquez, J. M. Guerrero, and F. Bullo, "Secondary frequency and voltage control of islanded microgrids via distributed averaging," *IEEE Trans. Ind. Electron.*, vol. 62, no. 11, pp. 7025–7038, 2015.

[39] S. Shah, H. Sun, D. Nikovski, and J. Zhang, "Consensus-based synchronization of microgrids at multiple points of interconnection," in *Proc. of the 2018 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE, 2018, pp. 1–5.

[40] Y. Wang, H. Ishii, F. Bonnet, and X. Défago, "Resilient consensus for multi-agent systems under adversarial spreading processes," *IEEE Trans. Netw. Sci. Eng.*, vol. 9, no. 5, pp. 3316–3331, 2022.

Zirui Liao received the B.E. degree in mechanical engineering from China Agricultural University, Beijing, China, in 2020. From 2022 to 2024, he was a visiting Ph. D. student at Eindhoven University of Technology, Eindhoven, The Netherlands.

He is currently pursuing the Ph.D. degree in mechanical engineering with Beihang University, Beijing, China. His research interests include cyber-physical system, resilient control, distributed optimization.

Jian Shi received the Ph.D. degree in mechanical engineering from Beihang University, Beijing, China, in 2007.

He is currently an Associate Professor with the School of Automation Science and Electrical Engineering, Beihang University. His major research interests include fault diagnosis, health management, and network reliability.

Yuwei Zhang received the B.S. degree in mathematics and applied mathematics and the Ph.D. degree in mechanical engineering from Beihang University, Beijing, China, in 2016 and 2021, respectively.

From 2019 to 2021, he was a Visiting Research Student with IRIDIA, the Artificial Intelligence Lab, Université Libre de Bruxelles, Brussels, Belgium. He is currently a Postdoctoral Researcher with the School of Automation Science and Electrical Engineering, Beihang University. His research interests include nonlinear control of marine vehicles and cooperative control of robot swarm.

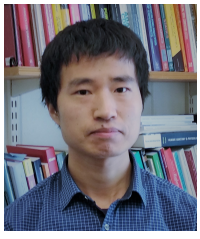
Shaoping Wang received the B.S., M.S., and Ph.D. degrees in mechanical engineering from Beihang University, Beijing, China, in 1988, 1991, and 1994, respectively.

She has been with the School of Automation Science and Electrical Engineering, Beihang University since 1994 and promoted to the Rank of Professor in 2000. Her research interests include engineering reliability, fault diagnostic, prognostic, health management, and fault-tolerant control.

Dr. Wang was honored as a Changjiang Scholar Professor by the Ministry of Education of China in 2013.

Rentong Chen received the B.S. degree from Northeastern University, Shenyang, China, in 2017. He received the Ph. D. degree in Mechatronics Engineering from Beihang University, Beijing, China, in 2023. From 2021 to 2022, he was a visiting Ph. D. student at Politecnico di Milano, Italy.

He is currently a postdoctoral researcher with the School of Automation Science and Electrical Engineering, Beihang University. His research interests include reliability theory, engineering and applied statistical learning.



Zhiyong Sun received the Ph.D. degree in control engineering from The Australian National University (ANU), Canberra, ACT, Australia, in February 2017.

He was a Research Fellow with ANU, and then a postdoctoral researcher at the Department of Automatic Control, Lund University of Sweden. Since January 2020, he has joined Eindhoven University of Technology (TU/e) as an assistant professor. He joins Peking University of China in the summer of 2024. His research interests include multi-agent systems, control of autonomous formations, distributed control and optimization.