Almost Sure Resilient Consensus Under Stochastic Interaction: Links Failure and Noisy Channels

Hamed Rezaee, *Member, IEEE*, Thomas Parisini, *Fellow, IEEE*, and Marios M. Polycarpou, *Fellow, IEEE*

***Abstract*— The resilient consensus problem over a class of discrete-time linear multiagent systems is addressed. Because of external cyber-attacks, some agents are assumed to be malicious and not following a desired cooperative behavior. Thus, the objective consists in designing a control strategy for the healthy agents to reach consensus upon their state vectors, while due to interaction among the agents, the malicious agents try to prevent them to achieve consensus. Although this problem has been investigated by some researchers, under the existing approaches in the literature, achieving consensus is only guaranteed when the information exchange among the agents is deterministic. Based on this motivation, the main contribution of the paper is on almost sure resilient consensus control of a network of healthy agents in the presence of stochastic links failure and communication noises. We design a discrete-time protocol for the set of the healthy agents, and we show that under some probabilistic conditions on interaction among the agents, achieving almost sure consensus among the healthy agents can be guaranteed. The results also are verified by numerical examples.**

***Index Terms*— Almost sure consensus, communication noises, cyber-physical systems, discrete-time, malicious agents, multiagent systems, stochastic topology.**

I. Introduction

COOPERATION in networks of autonomous agents has emerged as an important topic of research in various engineering areas such as space missions [1], [2], cooperative robots [3], [4], power networks [5], [6], etc. To accomplish a cooperative task, due to distributed behaviors of multiagent systems (MASs), it may be necessary for the agents to achieve consensus upon some quantities of interest via local interaction under a communication network [7]–[12]. One of important obstacles in achieving consensus in a network of agents is the malicious behavior of some unknown agents because of possible external cyber-attacks. The malicious agents do not follow the desired cooperative control strategy designed for the MAS. Thus, as they are unknown, they can deteriorate the performance of the MAS and may lead to divergence and instability in all the agents behaviors. The cyber-security of such systems is usually ensured by suitable information and communication technologies. However, these technologies may not be effective in all possible scenarios, and it is necessary to consider cyber-security in the MASs in the control layer by designing appropriate resilient control strategies [13]–[15].

A. State of the Art and Existing Problems

Primary studies on resilient consensus control of dynamical agents have been done on networks of single-integrator agents in [15] and [16]. The main idea of those studies is planning a strategy such that each agent detects and ignores any probable anomaly in its neighborhood. In [17] and [18], graph theoretic resilient consensus control strategies are presented in which sufficient conditions on the network robustness for achieving consensus are derived. The mentioned strategy is extended to achieve resilient consensus in a network of agents with double-integrator model in [19], and is extended to a network of agents with quantized values and randomized updating times in [20]. In [21], the performance of the mentioned strategy in the presence of communication delays is investigated. In [22], that idea is employed for resilient synchronization of high-order MASs. In [23] and [24], resilient consensus control of MASs with asynchronous update time is investigated. In [25], the resilient max consensus problem is studied. In [26], the resilient leader-follower control of MASs based on the idea of graph robustness is addressed. In [27], to relax the condition of graph robustness in a network of first-order agents, it is supposed that some agents are trustworthy for other agents and are not under cyber-attacks. In [28], a consensus protocol resilient against message manipulation attacks for time synchronization over sensor networks is proposed, and in [29], based on an expected/desired cooperative behavior from each agent, a reputation-based consensus control scheme for single-integrator MASs under cyber-attacks is proposed.

By investigation of the existing results on resilient consensus control of MASs, the following issues are noteworthy:

This work has been partially supported by European Union's Horizon 2020 research and innovation program under grant agreement no. 739551 (KIOS CoE) and by the Italian Ministry for Research in the framework of the 2017 Program for Research Projects of National Interest (PRIN), Grant no. 2017YKXYXJ.

H. Rezaee is with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, UK (e-mail: h.rezaee@imperial.ac.uk).

T. Parisini is with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, UK, with the Department of Engineering and Architecture, University of Trieste, 34127 Trieste, Italy, and also with the KIOS Research and Innovation Center of Excellence, University of Cyprus, CY-1678 Nicosia, Cyprus (e-mail: t.parisini@imperial.ac.uk).

M. M. Polycarpou is with the KIOS Research and Innovation Center of Excellence and the Department of Electrical and Computer Engineering, University of Cyprus, CY-1678 Nicosia, Cyprus (e-mail: mpolycar@ucy.ac.cy).

1) The base of the existing studies is deterministic distinguishing/selection of safer interaction links such that the malicious behaviors of agents are ignored over time. However, because of stochastic properties of devices or employing randomized communication protocols [30], the communication links may not be deterministic in many cases, and they may fail at any time instant. Thus, in practice, selection of safer communication links may not be deterministic.
2) The existing results are based on accurate communication among the agents and are not suitable for networks with noisy channels. Indeed, in the presence of communication noises, achieving consensus does not imply an equilibrium condition for the MAS as the measured consensus errors are noisy. Moreover, since communication noises lead to stochastic errors in received information, they may affect distinguishing/selection of safer state information received from neighboring agents.

B. Objectives and Contributions

In this paper, consensus control of a network of agents with stochastic interaction is dealt with. We consider a class of linear MASs in which some agents have malicious behaviors. In this condition, the objective is to address a resilient consensus control scheme under which the healthy agents reach consensus on a common state vector, despite the presence of the malicious agents that try to prevent the consensus achievement. In summary, the main contributions of this paper are:

1) Resilient consensus control of MASs when the communication links stochastically fail and rebuild over time.
2) Resilient consensus control of a network of agents when information exchange among the agents is not accurate because of stochastic noises.

To achieve these goals, we define a safety variable for each agent which will be shared with other agents through a stochastic network. Then, we design a control strategy under which each healthy agent at each time instant evaluates the stochastic safety variables of its neighbors and uses only the safest information to update its own states. Upon some probabilistic conditions on the noises and the communication graph robustness, achieving almost sure consensus among the healthy agents is addressed.

Remark 1: It is worth mentioning that the challenges of resilient control of MASs are different from those in fault tolerant control addressed in the literature [31]–[35]. Indeed, in the presence of cyber-attacks, an agent may be under the control of an attacker and be sagacious. Thus, a cyber-attack may not be modeled as a fault. However, since a faulty agent is a special form of a malicious agent, the obtained results for resilient control of MASs are also useful to filter out and evade faulty behaviors as well, such that a desired performance for the group of healthy agents can be guaranteed.

The organization of this paper is as follows. Preliminaries are provided in Section II. Motivation of the study is presented in Section III. In Section IV, the problem is stated. The proposed resilient consensus control strategy in the presence of stochastic links failure is presented in Section V. In Section VI, the results are extended to networks with noisy channels. Simulation results are presented in Section VII, and Section VIII concludes the paper.

II. Preliminaries

Notation and some concepts and definitions on graph theory and stochastic variables that are needed in the paper are provided in this section.

A. Notation

Let $\mathbb{R}$, $\mathbb{R}_{>0}$, and $\mathbb{R}_{\geq 0}$ be the sets of real, positive real, and nonnegative real numbers, respectively. $\mathbb{N}$ denotes the set of nonnegative integer numbers. 1_n is an $n \times 1$ ones vector. 0_n is an $n \times 1$ zeros vector. $\wedge$ denotes 'and' and $\vee$ denotes 'or'. $v(p(t))$ expresses a stochastic switching parameter where $p(t)$ (p in short) is the index associated with the switching set with n_v members such that $p(t) : [0, \infty) \to \{1, 2, \ldots, n_v\}$. $\mathbb{E}\{\cdot\}$ and $\mathbb{P}\{\cdot\}$ express the expected value and probability, respectively, and $\mathbb{E}\{X|E\}$ denotes the conditional expected value of X given an event E. For any scalar x, $|x|$ is the absolute value, and for a set $\mathcal{S}$, $|\mathcal{S}|$ denotes the cardinality. $\|\cdot\|$ denotes the absolute Euclidean norm. For two sets $\mathcal{S}_1$ and $\mathcal{S}_2$, $\mathcal{S}_1 \backslash \mathcal{S}_2$ denotes the reduction of $\mathcal{S}_1$ by $\mathcal{S}_2$. For two sets $\mathcal{S}_1$ and $\mathcal{S}_2$, we say $\mathcal{S}_1 \geq \mathcal{S}_2$ if any member of $\mathcal{S}_1$ is greater than or equal to any member of $\mathcal{S}_2$. $\det(\cdot)$ stands for the determinant. Moreover, '$\max$' means 'maximum', '$\min$' means 'minimum', 'a.s.' denotes 'almost surely', 'i.p.' denotes 'in the sense of probability', and 'w.p.' denotes 'with probability'.

B. Graph Theory

Interaction among the agents is described by a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$ where $\mathcal{V} = \{1, 2, \ldots, N\}$ denotes the set of N agents or nodes and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of communication links or edges where an edge (j, i) denotes that the ith agent receives information from the jth one. Under this condition, we say that the jth agent is a neighbor of the ith agent, and accordingly, $\mathcal{N}_i$ is defined as the neighboring set of the ith agent. Moreover, $\mathcal{A}$ is the adjacency matrix associated with $\mathcal{G}$ expressed as $\mathcal{A} = [a_{ij}] \in \mathbb{R}^{N \times N}$ where $a_{ij} \in \mathbb{R}_{>0}$ if $(j, i) \in \mathcal{E}, i \neq j$, and it is zero, otherwise. Indeed, the value of a_{ij} determines the weight of the edge from the jth node to the ith node. Such weights are used in the consensus protocols introduced in Sections V and VI.

For a directed graph $\mathcal{G}$, a nonempty set $\mathcal{S} \subset \mathcal{V}$ is said to be r-*reachable* if $\exists i \in \mathcal{S}$ s.t. $|\mathcal{N}_i \backslash \mathcal{S}| \geq r$. According to this definition, a directed graph $\mathcal{G}$ is said to be r-*robust* if for each two disjoint nonempty sets $\mathcal{S}_1, \mathcal{S}_2 \subset \mathcal{V}$ such that $\mathcal{S}_1 \cap \mathcal{S}_2 = \varnothing$, at least one of them is r-reachable. Moreover, a set of nodes $\mathcal{S} \subset \mathcal{V}$ is called f-*local* if for $i \in \mathcal{V} \backslash \mathcal{S}$, $|\mathcal{N}_i \cap \mathcal{S}| \leq f$ [17].

C. Stochastic Processes

The stochastic behavior of a process is described by the triple $(\Omega, \mathcal{F}, \mathbb{P})$ where Ω is the sample space, $\mathcal{F}$ is a σ-algebra on Ω, and $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{F})$ where $0 \leq \mathbb{P}\{\cdot\} \leq 1$ and $\mathbb{P}\{\Omega\} = 1$ [36].

A filtration $\{\mathcal{F}_r, r \geq 0\}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ is a sequence of sub σ-algebras of $\mathcal{F}$ such that $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \ldots \subseteq \mathcal{F}$. Based on the definition of filtration, a stochastic process $X = \{X(r), r \geq 0\}$ is said to be adapted to the filtration $\{\mathcal{F}_r\}$ if for each r, $X(r)$ is $\mathcal{F}_r$-measurable. Now, based on the definition of adapted stochastic processes, a process X is a *super-martingale* relative to $\{\mathcal{F}_r\}$ and $\mathbb{P}$ if [36, Chap. 10]:

i) X is adapted to the filtration $\{\mathcal{F}_r\}$,
ii) $\mathbb{E}\{|X(r)|\} < \infty, \forall r$,
iii) $\mathbb{E}\{X(r)|\mathcal{F}_{r-1}\} \leq X(r-1), r \geq 1$.

Moreover, a process X is a *martingale difference sequence (MDS)* relative to $\{\mathcal{F}_r\}$ and $\mathbb{P}$ if the above third condition is changed as follows [37]:

$$\mathbb{E}\{X(r)|\mathcal{F}_{r-1}\} = 0, r \geq 1.$$

For instance, a white noise is a special form of MDSs.

The stochastic variable $X(t)$ converges to X_f *in the sense of probability* if [38]

$$\lim_{t\to\infty} \mathbb{P}\{|X(t) - X_f| \geq \epsilon\} = 0, \forall \epsilon \in \mathbb{R}_{>0}.$$

In this case, we write $\lim_{t\to\infty} X(t) \xrightarrow{\text{i.p.}} X_f$. Moreover, the stochastic variable $X(t)$ converges to X_f *almost surely* if [38]

$$\mathbb{P}\{\lim_{t\to\infty} X(t) = X_f\} = 1.$$

In this case, we write $\lim_{t\to\infty} X(t) \xrightarrow{\text{a.s.}} X_f$. In general, we say that an event *occurs almost surely*, if the probability of other events is zero. It is worth mentioning that convergence in the sense of probability is the weaker criterion in which nonzero errors with zero probability still may happen.

III. Motivation

Achieving consensus in a MAS is based on information exchange among the agents via a sufficiently connected network such that each agent is an attraction point for a group of agents, while it is attracted toward neighboring agents [7]. The attraction of a team of two-dimensional agents toward each other is shown in Fig. 1. According to the figure, the distance between the maximum and minimum values of the agents states (in each dimension) is decreased over time such that consensus is achieved in the MAS. However, decreasing of such distances may not be realizable if some agents are malicious. For instance, consider a case when one of the agents has an unstable trajectory. As shown in Fig. 2, if this malicious agent, shown by a red circle, is unknown to the other agents, it can attract them toward itself and can lead to divergence in the MAS.

To cope with this issue, the concept of resilient consensus has been developed. The main idea of the resilient consensus problem is that while each agent updates its states toward the states values of its neighbors, it ignores some of them based on the knowledge of the maximum number of possible malicious neighbors. In [17] and [18], it was analyzed that to keep consensusability after ignoring some neighbors, more communication links among the agents should be established.

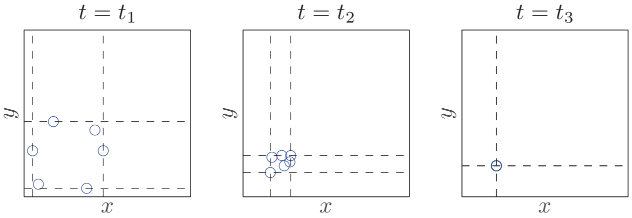

Fig. 1. Convergence of the maximum value and the minimum value of six agents states (in two dimensions) toward each other to achieve consensus.

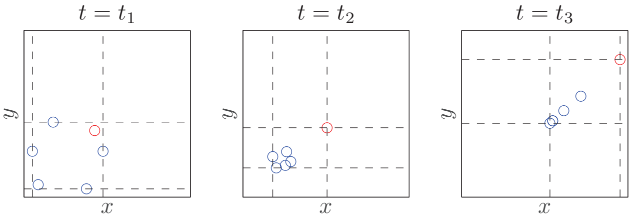

Fig. 2. Divergence of the states values of six agents when one of the agents (the red circle) does not follow the consensus strategy, while according to the consensus strategy, it attracts the healthy agents toward itself.

However, the established links may still not be available constantly, and they may fail over time randomly. In this condition, the evaluation of the states of the neighboring agents and selecting safer ones is a random process. Hence, the design and analysis tools available in the literature for resilient consensus control are not applicable.

Another important practical concern not considered in the literature of the resilient consensus problem is that information exchange among the agents may not be precise, because of noisy communication channels. In the presence of communication noises, the steady state of the MAS may be noisy, and zero consensus errors may not imply an equilibrium condition for the MAS. Moreover, when the channels are noisy, received state information from neighboring agents may jump to incorrect values at some time instants. In this condition, each healthy agent should evaluate, select, and use states which however are erroneous due to communication noises.

The consensus problem in the presence of stochastic links failure and noisy channels is already considered in the literature [39]–[44]. However, when the resiliency property is considered, the evaluation and selection of safer neighboring agents by each healthy agent has to be considered, and this leads to “state-dependent switching laws” that are not dealt with in those studies. Thus, in such case, the existing control strategies and analysis tools for stochastic networks are not applicable to guarantee the achievement of consensus in the MAS. Moreover, achieving resiliency in the sense mentioned above is even more challenging in the presence of noisy channels, since the mentioned state-dependent switching law is also based on noisy and imprecise stochastic information.

Based on the above-mentioned issues, it is worth proposing and analyzing a resilient consensus strategy such that when interaction among the agents is not reliable, achieving consensus in the network is still achievable.

IV. Problem Statement

Consider a class of discrete-time MASs comprising of N agents where the model of the ith agent is described by

$$x_i(t+1) = Ax_i(t) + Bu_i(t) \tag{1}$$

where $x_i(t) \in \mathbb{R}^n$ is the state vector which is assumed to be measurable, $A \in \mathbb{R}^{n\times n}$ is the state matrix, $B \in \mathbb{R}^{n\times 1}$ is the input matrix, and $u_i(t) \in \mathbb{R}$ is the control input.

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

As the system is controllable, we can consider a similarity transformation $\xi_i(t) = T^{-1}x_i(t)$ where if

$$\det(zI_n - A) = z^n + \gamma_1 z^{n-1} + \gamma_2 z^{n-2} + \ldots + \gamma_n, \tag{2}$$

the MAS described in (1) can be described by the following companion model [45]:

$$\xi_i(t+1) = \bar{A}\xi_i(t) + \bar{B}u_i(t) \tag{3}$$

in which

$$\bar{A} = \begin{bmatrix} -\gamma_1 & -\gamma_2 & \ldots & -\gamma_{n-1} & -\gamma_n \\ 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \end{bmatrix}, \bar{B} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

This transformation is used in presentation of the main results.

Because of cyber-attacks, some agents are assumed to have malicious behaviors. Indeed, an agent in the MAS is called malicious (also called Byzantine [16], [18]) if it has at least one of the following behaviors:

- It updates its states in any way other than what is designed/desired in the consensus protocol.
- It shares fake/wrong state information in the network other than its real state information (the shared information in various communication links may be different).

Note that this definition of attacks covers a wide range of cyber-attacks in practice few of which are as follows:

- *Data injection attacks*: They are a type of attacks in network systems when the information flow from a subsystem to another one is corrupted by injecting some undesirable data. For instance, in the case of data injection attacks in communicated state information, by defining $x^c_{ij}(t)$ as the corrupted state information of the jth agent received by the ith agent, it can be described as [12]

  $$x^c_{ij}(t) = x_j(t) + \beta_{ij}(t)\delta_{ij}(t)$$

  where $\beta_{ij}(t) \in \{0, 1\}$ denotes the attack activation function and $\delta_{ij}(t)$ is an unknown data injected by the attacker. Thus, the jth agent shares fake/wrong state information with the ith one. Based on a similar argument, one can say that in the case of such attacks in the actuators or sensors of an agent, the agent updates its states in a way other than what is designed/desired. Sensor attacks lead to sharing fake/wrong state information as well.
- *Replay attacks*: Replay attacks happen when an attacker records transmitted information over a network and then replays/repeats it instead of the real information. For instance, in the case of such attacks in communicated state information, if we define $x^r_{ij}(t)$ as the replayed information of the jth agent received by the ith agent, it can be modeled as [46]

  $$x^r_{ij}(t) = x_j(t) + \beta_{ij}(t)\big(-x_j(t) + x_j(t - T_{ij}(t))\big)$$

  where $T_{ij}(t) \in \mathbb{R}_{\geq 0}$ denotes a time-delay. Thus, replaying communicated state information leads to sharing fake/wrong state information in the network.
- *Denial of service attacks*: Under a denial of service attack, the information flow between two components will be prevented. For instance, in the case of such attacks between the control unit and the actuators of an agent, the real control command affecting the agent can be [47]

  $$u^a_i(t) = (1 - \beta_i(t))u_i(t)$$

  where $\beta_i(t) \in \{0, 1\}$ denotes the attack activation function. In this condition, the agent may not update its states in a desired way.

Thus, in the presence of such cyber-attacks, we have two groups of agents: N_h healthy agents described by the fixed set $\mathcal{V}_h$ and N_m unknown malicious agents described by the fixed set $\mathcal{V}_m$ where $N_h + N_m = N$ and $\mathcal{V}_h \cup \mathcal{V}_m = \mathcal{V}$. Since the malicious agents are under the control or partial control of the attacker, the objective is designing an interaction consensus protocol for the healthy agents such that while the malicious agents try to prevent them to achieve consensus, they reach safe consensus upon their state vectors as

$$\lim_{t\to\infty}(x_i(t) - x_j(t)) = 0_n, i, j \in \mathcal{V}_h, \tag{4}$$

and all the states remain bounded (with a bound depending on initial states) during transient times.

Assumption 2: While the malicious agents are considered unknown to the healthy agents, we assume that the set of the malicious agents is f-local where f is known, i.e., we assume that the worst case of the number of malicious agents in the neighborhood of each healthy agent is known.

Remark 2: Throughout the paper, any control strategy and decision making scheme is planned for the healthy agents as we have no control or partial control on the malicious agents. Indeed, the details of the malicious agents behaviors are not relevant, and the objective is to propose a consensus control strategy that acts while being resilient against malicious behaviors of some agents (due to any source/type of cyber-attacks).

In order to achieve consensus among the healthy agents, the agents should share their state information with each other through communication links. Due to random availability of these communication links, their connectivity is described in probabilistic terms: the connectivity of the link from Agent j to Agent i, where $i, j \in \mathcal{V}$, is modeled by the time-varying probability $p_{ij}(t) \in [0, 1]$ (if $\forall t$ there is no a communication link from Agent j to Agent i, we have $p_{ij}(t) = 0, \forall t$). In this condition, the adjacency matrix of the communication graph will be stochastic defined as $\mathcal{A}(p) = [a_{ij}(p)]$ where $a_{ij}(p)$ is a stochastic switching parameter with the following stochastic switching law:

$$a_{ij}(p) = \begin{cases} \in \mathbb{R}_{>0} & \text{w.p.} \quad p_{ij}(t) \\ 0 & \text{w.p.} \quad 1 - p_{ij}(t). \end{cases}$$

Accordingly, the communication graph will be stated as $\mathcal{G}(p) = (\mathcal{V}, \mathcal{E}(p), \mathcal{A}(p))$ and the ith agent neighborhood will be $\mathcal{N}_i(p)$. Due to the stochastic properties of the closed-loop system, it is necessary to study achieving consensus among the healthy agents in the view of stochastic processes theory. Thus, our objective is achieving almost sure safe consensus in the network such that the objective (4) is modified as

$$\lim_{t\to\infty} (x_i(t) - x_j(t)) \xrightarrow{\text{a.s.}} 0_n, i, j \in \mathcal{V}_h, \tag{5}$$

while all the states almost surely remain bounded (with an expected bound depending on initial states) during transient times.

The main results are presented in the subsequent sections.

V. Resilient Consensus Under Stochastic Links Failure

The basic idea in resilient consensus control of MASs is the evaluation of the state information of neighboring agents and ignoring some of them in updating own states. Thus, based on agents models and existing practical issues, designing a resilient consensus control strategy for a network of agents has two main challenges [15]–[19]:

1) How should each healthy agent distinguish and select neighbors which do not lead to divergence?
2) How should it employ the selected neighbors state information to achieve safe consensus with other healthy agents?

To develop a resilient consensus strategy in the presence of stochastic links failure, we consider a criterion to guarantee safety in the network under which the states of the healthy agents almost surely remain in a bound. According to this criterion, the ith agent shares a safety variable $s_i(t)$ with other agents, and the value of this safety variable will be investigated by the jth healthy agent to choose this agent for interaction or not (if $i \in \mathcal{N}_j(p), j \in \mathcal{V}_h$). Then, if the ith agent is chosen by a healthy agent for interaction, $s_i(t)$ will be used by that healthy agent in an interaction consensus protocol such that consensus is achieved in the network eventually. Indeed, each healthy agent at each time instant updates its states just based on the safety variables of those neighbors which do not lead to divergence of its own safety variable (this safety variable should be designed such that its boundedness guarantees the boundedness of all the states of the agent, which is designed later). Therefore, we have two reasons for links cutting in the network:

- The primary is stochastic failure of communication links discussed in the previous section.
- The secondary is deliberate ignoring of some links/neighbors by each healthy agent based on a designed resilient consensus strategy. In this case, we define a variable $k_{ij}(t), i \in \mathcal{V}_h, j \in \mathcal{N}_i(p)$, where $k_{ij}(t) = 1$ describes that the ith agent selects the jth neighbor; otherwise, $k_{ij}(t) = 0$.

Therefore, after the selection/ignoring strategy, we define the network *effective adjacency matrix* as $\mathcal{A}(p,t) = [a_{ij}(p,t)]$ where

$$a_{ij}(p,t) = \begin{cases} \in \mathbb{R}_{>0} & \text{w.p.} \;\; p_{ij}(t) \quad\;\; \wedge \quad k_{ij}(t) = 1 \\ 0 & \text{w.p.} \;\; 1 - p_{ij}(t) \;\; \vee \quad k_{ij}(t) = 0. \end{cases}$$

In a similar way, the *effective neighboring set* of each agent will be defined as $\mathcal{N}_i(p,t)$. Note that we only deal with $a_{ij}(p,t)$ if $i \in \mathcal{V}_h$, because the malicious agents are not under our control, and thus the entries of the adjacency matrix if $i \in \mathcal{V}_m$ are not important for us.

After selecting the safest neighbors, a proper control strategy should be employed such that guarantees reaching consensus among the healthy agents, while it guarantees the boundedness of the healthy agents safety variables. Therefore, we propose our resilient consensus control strategy for the healthy agents in **two parts**:

(a) The healthy agent i receives the information $s_j(t)$ if $j \in \mathcal{N}_i(p)$, and sets $k_{ij}(t) = 1, j \in \mathcal{N}_i(p)$. If $|\mathcal{N}_i(p)| \geq f$, it considers f neighbors with largest $s_j(t)$ and if $|\mathcal{N}_i(p)| < f$, it considers all the $|\mathcal{N}_i(p)|$ neighbors. Then, for a neighbor j if $s_j(t) > s_i(t)$, it sets $k_{ij}(t) = 0$. In a similar way, if $|\mathcal{N}_i(p)| \geq f$, it considers f neighbors with smallest $s_j(t)$ and if $|\mathcal{N}_i(p)| < f$, it considers all the $|\mathcal{N}_i(p)|$ neighbors. Then, for a neighbor j, if $s_j(t) < s_i(t)$, it sets $k_{ij}(t) = 0$ (the idea of selection based on the knowledge of f is inspired by the existing literature for first-order systems [15]–[18]). Since the boundedness of $s_i(t)$ should guarantee the boundedness of the healthy agents states, by stating $\xi_i(t)$ as

$$\xi_i(t) = \begin{bmatrix} \xi_{n-1,i}(t) & \xi_{n-2,i}(t) & \ldots & \xi_{0,i}(t) \end{bmatrix}^\top, \tag{6}$$

we have proposed it as follows:

$$s_i(t) = \xi_{n-1,i}(t) + \sum_{m=1}^{n-1} \lambda_m \xi_{n-1-m,i}(t), i \in \mathcal{V}_h, \tag{7}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_{n-1} \in \mathbb{R}$ are chosen such that the following polynomial is Schur stable:

$$\text{pol}(z) = z^{n-1} + \lambda_1 z^{n-2} + \lambda_2 z^{n-3} + \ldots + \lambda_{n-1}. \tag{8}$$

The details of the boundedness of the healthy agents states under the boundedness of the safety variables will be discussed later.

(b) We should design a consensus strategy such that by using the state information of the selected stochastic neighbors in Part (a) and by considering the criterion which has selected these neighbors, the healthy agents reach safe almost sure consensus upon their state vectors.

Based on the mentioned consensus control strategy in Parts (a) and (b), the remained and challenging issue is designing u_i in Part (b) such that the requirements of Parts (a) and (b) are satisfied.

The proposed resilient control strategy for MASs in the presence of stochastic links failure is proposed in the following theorem.

Definition 1: For a stochastic graph $\mathcal{G}(p)$, let $\mathcal{E}_c(t)$ be the set of edges which are probable to be connected at time t. In other words, let $\mathcal{E}_c(t) = \{(j,i)|p_{ij}(t) \neq 0\}$. We say that the set of the stochastic edges of the graph $\mathcal{G}(p)$ with nonzero probability makes the graph k-robust, if $\mathcal{E}_c(t)$ always makes the graph k-robust. Then, we write $\mathbb{P}\{\mathcal{G}(p)$ is k-robust$\} \neq 0$.

It is worth noting that a k-robust graph is a special form of a graph with the condition $\mathbb{P}\{\mathcal{G}(p)$ is k-robust$\} \neq 0$. Indeed, a graph with the condition $\mathbb{P}\{\mathcal{G}(p)$ is k-robust$\} \neq 0$ can be obtained by considering a k-robust graph and by letting the edges fail and rebuild over time (with nonzero probability of connectivity). Thus, the condition $\mathbb{P}\{\mathcal{G}(p)$ is k-robust$\} \neq 0$ is weaker than the condition of being k-robust.

Theorem 1: Consider a network of N agents described in (1) containing N_h healthy agents and N_m malicious agents. Let the malicious agents set be f-local and $\mathbb{P}\{\mathcal{G}(p)$ is $(2f+1)$-robust$\} \neq 0$. Moreover, the safety variables are chosen as (7), and accordingly the gains $k_{ij}(t), i \in \mathcal{V}_h, j \in \mathcal{N}_i(p)$, follow the proposed selection strategy in Part (a). The following consensus protocol guarantees almost sure safe consensus among the healthy agents:

$$\begin{aligned} u_i = & \sum_{m=1}^{n} \gamma_m \xi_{n-m,i}(t) - \sum_{m=1}^{n-1} \lambda_m \xi_{n-m,i}(t) + s_i(t) \\ & + \alpha_i(p,t) \sum_{j=1}^{N} a_{ij}(p,t) \big[s_j(t) - s_i(t) \big], i \in \mathcal{V}_h, \end{aligned} \tag{9}$$

in which at each time instant, $\alpha_i(p,t)$ should be chosen such that for an arbitrary $\theta_i \in \mathbb{R}_{>0}$,

$$0 < \alpha_i(p,t) \leq \frac{1}{\theta_i + \sum_{j=1}^{N} a_{ij}(p,t)}. \tag{10}$$

Proof: See the appendix for the proof. ■

By considering all the mentioned issues in this section, the resilient consensus control algorithm of Theorem 1 at step time t is summarized in Algorithm 1.

Algorithm 1 Consensus algorithm of Theorem 1 at time t

```
1: The healthy agent $i$ receives the information $s_j(t)$ if $j \in \mathcal{N}_i(p)$, and sets $k_{ij}(t) = 1, j \in \mathcal{N}_i(p)$.
2: If $|\mathcal{N}_i(p)| \geq f$, it considers $f$ stochastic neighbors with largest $s_j(t)$ and if $|\mathcal{N}_i(p)| < f$, it considers all the $|\mathcal{N}_i(p)|$ stochastic neighbors. Then, if $s_j(t) > s_i(t)$, it sets $k_{ij}(t) = 0$.
3: If $|\mathcal{N}_i(p)| \geq f$, it considers $f$ stochastic neighbors with smallest $s_j(t)$ and if $|\mathcal{N}_i(p)| < f$, it considers all the $|\mathcal{N}_i(p)|$ stochastic neighbors. Then, if $s_j(t) < s_i(t)$, it sets $k_{ij}(t) = 0$.
4: By calculating the effective adjacency matrix entries $a_{ij}(p,t)$ from $k_{ij}(t)$, the agent updates its states via the proposed interaction protocol (9).
```

Remark 3: It is noteworthy that Part (a) of the consensus control strategy does not imply that all the ignored neighbors are malicious or all the malicious neighbors are ignored. Indeed, based on the evaluation of safety variables, only the neighbors with the safest behaviors will be selected by each healthy agent. Therefore, the cooperation of a malicious agent may be safe at some time instants, if the transmitted value of its safety variable is inside a range such that it is not ignored in Part (a) of the consensus control strategy (no matter the transmitted value is fake or real).

Remark 4: In the consensus protocol (9), for the case of $n = 1$, all the summations from the index 1 to $n-1$ should be considered zero.

The obtained results will be extended to noisy networks of agents in the next section.

VI. Resilient Consensus Under Stochastic Links Failure and Noisy Channels

In this section, we consider the problem of resilient consensus under stochastic links failure when the agents are also prone to stochastic noises in communication. We assume that each agent receives the information of its neighbors via a channel subject to noises with MDS properties independent to stochastic failure of the links. In this condition, if the jth agent sends $s_j(t)$ to the ith agent, the ith agent receives $\tilde{s}_{ij}(t)$ defined as follows:

$$\tilde{s}_{ij}(t) = s_j(t) + \omega_{s_{ij}}(t)$$

where $\omega_{s_{ij}}(t)$ describes the noise associated with communication of $s_j(t)$ to the ith agent. It should be noted that the main idea of the proposed control strategy in this section is using the MDS properties of the communication noises to filter their effects on the steady state of the MAS. Without this property, the communicated information can be similar to information sent by a malicious agent.

In the resilient consensus control strategy proposed in the previous section, communication noises in two ways can affect the performance of the healthy agents in achieving consensus. Firstly, in Part (a) of the consensus control strategy, the healthy agent i selects/ignores some of its neighbors by receiving and evaluating the noisy information of safety variables of its neighbors defined as $\tilde{s}_{ij}(t), j \in \mathcal{N}_i(p)$. Thus, the selection/ignoring strategy is prone to some stochastic errors. In this case, the selection parameter $k_{ij}(t), i \in \mathcal{V}_h, j \in \mathcal{N}_i(p)$, should be modified by $k_{ij}(t,\omega), i \in \mathcal{V}_h, j \in \mathcal{N}_i(p)$, to show that they are affected by noises. Moreover, we should modify the network effective adjacency matrix as $\mathcal{A}(p,t,\omega) = [a_{ij}(p,t,\omega)]$ where

$$a_{ij}(p,t,\omega) = \begin{cases} \mathbb{R}_{>0} & \text{w.p. } p_{ij}(t) \quad \wedge \quad k_{ij}(t,\omega) = 1 \\ 0 & \text{w.p. } 1 - p_{ij}(t) \quad \vee \quad k_{ij}(t,\omega) = 0, \end{cases}$$

and in a similar way, the effective neighboring set of each agent should be modified as $\mathcal{N}_i(p,t,\omega)$. The second way on which communication noises affect the performance of the healthy agents is that the consensus protocol designed in Part (b) of the consensus control strategy in Theorem 1 is based on the state information of some selected neighboring agents, while this information contains stochastic noises. Thus, by employing the consensus strategy proposed in Theorem 1, almost sure consensus may not be guaranteed. To cope with the mentioned problems, the consensus protocol of Theorem 1 will be modified and analyzed in the presence of communication noises. The main results are presented in a theorem as follows.

Theorem 2: Consider the MAS described in (1) containing N_h healthy agents and N_m malicious agents under a noisy communication network which the noises satisfy the conditions of MDSs. Let the malicious set be f-local,

$\mathbb{P}\{\mathcal{G}(p)$ is $(2f+1)$-robust$\} \neq 0$, the safety variables are chosen as (7), and the gains $k_{ij}(t,\omega), i \in \mathcal{V}_h, j \in \mathcal{N}_i(p)$, follow the proposed selection strategy in Part (a). The following consensus protocol guarantees almost sure safe consensus among the healthy agents:

$$\begin{aligned} u_i(t) &= \sum_{m=1}^{n} \gamma_m \xi_{n-m,i}(t) - \sum_{m=1}^{n-1} \lambda_m \xi_{n-m,i}(t) + s_i(t) \\ &\quad + \alpha_i(p,t,\omega) \sum_{j=1}^{N} a_{ij}(p,t,\omega)\big[\tilde{s}_{ij}(t) - s_i(t)\big], \qquad (11) \\ &\quad i \in \mathcal{V}_h, \end{aligned}$$

if at each time instant, $\alpha_i(p,t,\omega)$ is chosen such that for an arbitrary $\theta_i \in \mathbb{R}_{>0}$, the following properties hold:

$$0 < \alpha_i(p,t,\omega) \leq \frac{1}{\theta_i + \sum_{j=1}^{N} a_{ij}(p,t,\omega)}, \qquad (12a)$$

$$\lim_{t\to\infty} \alpha_i(p,t,\omega) = 0, \qquad (12b)$$

$$\sum_{t=t_0}^{\infty} \alpha_i(p,t,\omega) = \infty, \forall t_0 \geq 0. \qquad (12c)$$

Proof: See the appendix for the proof. ■

By considering all the mentioned issues in this section, the resilient consensus control algorithm of Theorem 2 at step time t is summarized in Algorithm 2.

Algorithm 2 Consensus algorithm of Theorem 2 at time t

```
1: The healthy agent $i$ receives the noisy information $\tilde{s}_{ij}(t)$
   if $j \in \mathcal{N}_i(p)$, and sets $k_{ij}(t,\omega) = 1, j \in \mathcal{N}_i(p)$.
2: If $|\mathcal{N}_i(p)| \geq f$, it considers $f$ stochastic neighbors with
   largest $\tilde{s}_{ij}(t)$ and if $|\mathcal{N}_i(p)| < f$, it considers all the
   $|\mathcal{N}_i(p)|$ stochastic neighbors. Then, if $\tilde{s}_{ij}(t) > s_i(t)$, it
   sets $k_{ij}(t,\omega) = 0$.
3: If $|\mathcal{N}_i(p)| \geq f$, it considers $f$ stochastic neighbors with
   smallest $\tilde{s}_{ij}(t)$ and if $|\mathcal{N}_i(p)| < f$, it considers all the
   $|\mathcal{N}_i(p)|$ stochastic neighbors. Then, if $\tilde{s}_{ij}(t) < s_i(t)$, it
   sets $k_{ij}(t,\omega) = 0$.
4: By calculating the effective adjacency matrix entries
   $a_{ij}(p,t,\omega)$ from $k_{ij}(t,\omega)$, the agent updates its states via
   the proposed interaction protocol (11).
```

It should be noted that the idea of using vanishing gains similar to (12) is also employed in the literature for consensus control of MASs in the presence of communication noises (for instance, see [43] and [44]). However, when resiliency property is considered, the existing design and analysis tools are not applicable to guarantee achieving consensus in the network.

Remark 5: Under the control strategies proposed in Theorem 1 and Theorem 2, if the number of malicious neighbors is not more than f, the healthy agents almost surely reach consensus upon their state vectors. According to these results, for cases when some agents are malicious only in finite time periods, the following issues are worthy to be noted:

- Because of finite time malicious behaviors of some agents, the number of malicious neighbors may be more than f in some finite periods of time, but in $t \geq t_f$ for a finite t_f, it is not more than f. Now, if the safety variable information transmitted by each malicious agent is bounded for $t < t_f$, according to (9) and (11), the healthy agents safety variables remain bounded for $t < t_f$ as well. In this condition, for $t \geq t_f$, since the number of the malicious neighbors is not more than f, if the safety variables of new added healthy agents are bounded at $t = t_f$, the control strategies proposed in Theorem 1 and Theorem 2 guarantee achieving almost sure consensus among the new fixed set of healthy agents.
- The number of malicious neighbors is not more than f in all time, but Agent i is malicious in some finite periods of time and it is healthy in $t \geq t_f$ for a finite t_f. In this condition, if $s_i(t_f)$ is bounded; then, for $t \geq t_f$, it behaves the same as a healthy agent and satisfies the results of Theorem 1 and Theorem 2 for the healthy agents.

Remark 6: It should be noted that Theorem 2 is applicable for both noisy and noise-free networks. However, because of using the property (12b), Theorem 2 leads to more conservative results compared with Theorem 1 as the property (12b) affects the transient response of the MAS. Therefore, for a noise-free network, the control strategy proposed in Theorem 1 is preferred to guarantee achieving consensus.

The proposed consensus control strategies will be evaluated via numerical examples in the following section.

VII. Numerical Examples

We consider the depth consensus problem in a network of twelve autonomous underwater vehicles (AUVs). The mathematical model of the depth dynamics of the ith vehicle is described as follows [48] (we have used the discrete-time model of the depth dynamics of an AUV with step size 1):

$$\begin{bmatrix} x_{1i}(t+1) \\ x_{2i}(t+1) \\ x_{3i}(t+1) \end{bmatrix} = \begin{bmatrix} 0.4037 & -0.2052 & 0 \\ 0.684 & 0.8825 & 0 \\ -0.1175 & -0.2875 & 1 \end{bmatrix} \begin{bmatrix} x_{1i}(t) \\ x_{2i}(t) \\ x_{3i}(t) \end{bmatrix} + \begin{bmatrix} 0.02394 \\ 0.01371 \\ -0.00146 \end{bmatrix} u_i(t)$$

where $x_{1i}(t)$ is the rate of the pitch angle, $x_{2i}(t)$ is the pitch angle, $x_{3i}(t)$ denotes the depth, and $u_i(t)$ is the control input. Moreover, from (2), it follows that $\gamma_1 = -2.2862$, $\gamma_2 = 1.7828$, and $\gamma_3 = -0.4966$, and therefore the similarity transformation matrix T can be obtained. The 6th and the 12th agents are supposed to be malicious because of external cyber-attacks, while the other agents are healthy and follow the desired consensus strategy proposed in Theorems 1 or 2. Without loss of generality, let the set of the malicious agents be 1-local. Thus, to guarantee that $\mathbb{P}\{\mathcal{G}(p)$ is 3-robust$\} \neq 0$, a communication graph as shown in Fig. 3 is considered, whose links are stochastic and the probability of the connectivity of each link is a time-varying number belonging to $[0.6, 0.8]$ (we have used various sinusoidal functions to model $p_{ij}(t), i, j \in \mathcal{V}$). The agents initial conditions are set arbitrarily and the consensus protocol gains are chosen as $\lambda_1 = 1/5$ and $\lambda_2 = 1/100$ to make the polynomial $z^2 + \lambda_1 z + \lambda_2$ Schur

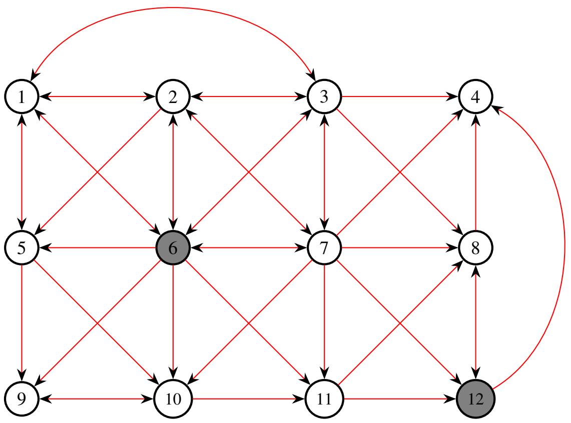

Fig. 3. Network communication graph when Agents 6 and 12 are malicious (each link stochastically fails over time).

stable. Moreover, for (9) and (11), at each time instant we set $a_{ij}(\cdot) = 1, i \in \mathcal{V}_h$, if the ith healthy agent uses the information of the jth one. In the following, three scenarios of malicious behaviors are considered.

Scenario 1 (Replay attacks in actuators): In the first scenario, we employ the proposed resilient consensus control strategy in Theorem 1 when the communication channels are assumed not to be noisy and just stochastic links failure among the agents in considered. Thus, to satisfy the condition of the theorem on $\alpha_i(p,t)$, let

$$\alpha_i(p,t) = \frac{1}{5\big(1+\sum_{j=1}^{N} a_{ij}(p,t)\big)}.$$

Moreover, the malicious agents are assumed to be subject to replay attacks in their actuators for $t \geq 20$s. Accordingly, without loss of generality, we model the replay attacks as 2s time delays in the actuators.

In this condition, as depicted in Fig. 4, whereas the malicious agents have unstable behaviors and are in interaction with the healthy agents, almost sure consensus among the healthy agents is achieved.

Scenario 2 (Simultaneous data injection and denial of service attacks in actuators and data injection attacks on communication links): In the second scenario, we consider the same control strategy as that in Scenario 1 when the communication channels are noisy. In this case, the malicious agents are assumed to be subject to simultaneous denial of service attacks and data injection attacks in their actuators for $t \geq 20$s, and also subject to data injection attacks in communicated information to other agents for $t \geq 25$s such that the injected data to various links are different (we have considered various sinusoidal injected signals with magnitudes belonging to $[-10, 10]$ to model the data injection attacks).

For simulation, we have employed the white noise tool in MATLAB with various powers to model the noisy channels. In this condition, the state trajectories of the agents are depicted in Fig. 5 showing that the performance is significantly deteriorated by the noises. Therefore, the resilient consensus control strategy of Theorem 1 is not effective in coping with the noises in the communication channels.

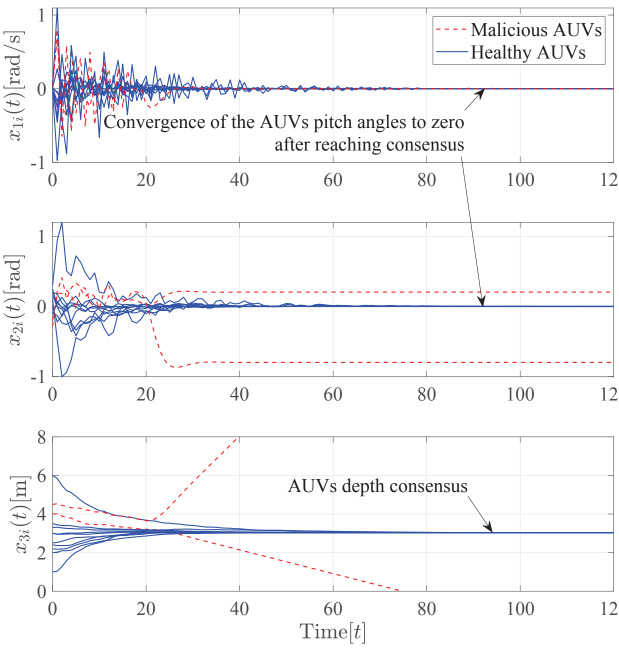

Fig. 4. State trajectories of the agents under the consensus protocol of Theorem 1 when the communication links fail stochastically.

Scenario 3 (Simultaneous data injection and denial of service attacks in actuators and data injection attacks on communication links): Finally, we consider Scenario 2 when the agents are under the action of the resilient consensus control strategy proposed in Theorem 2. In this case, to satisfy the conditions of the theorem, we design $\alpha_i(p,t,\omega)$ as

$$\alpha_i(p,t,\omega) = \frac{1}{(0.2t+4)\big(1+\sum_{j=1}^{N} a_{ij}(p,t,\omega)\big)}. \qquad (13)$$

Since $1/(0.2t+4) < 1$ and $\alpha_i(p,t,\omega)$ eventually converges to zero, the conditions (12a) and (12b) are satisfied. Moreover, by considering (13), it can be said that

$$\alpha_i(p,t,\omega) \geq \frac{1}{N(0.2t+4)}$$

implying that the condition (12c) is satisfied as well. As depicted in Fig. 6, the proposed strategy can cope with the problem of noisy channels and guarantee almost sure consensus in the network. It should be noted that, although the proposed strategy in Theorem 2 can also guarantee almost sure consensus under stochastic links failure without noisy channels, the results of this theorem are more conservative compared with those of Theorem 1. In fact, since $1/(0.2t+4)$ is asymptotically vanishing; then, the convergence time for achieving consensus is increased (see Fig. 6).

VIII. Conclusions and Future Work

In this study, a consensus control framework for a class of discrete-time linear MASs in the presence of cyber-attacks was proposed. We developed a resilient control strategy under which a group of healthy agents tried to employ the safest available information exchanged via the network, while some malicious agents tried to prevent them to achieve consensus.

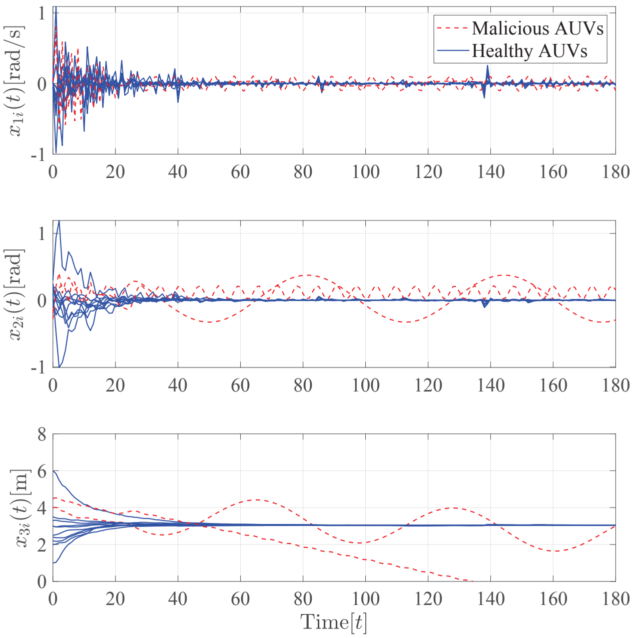

Fig. 5. Deterioration of the performance of the consensus protocol of Theorem 1 when the communication links fail stochastically and the communication channels are subject to noises.

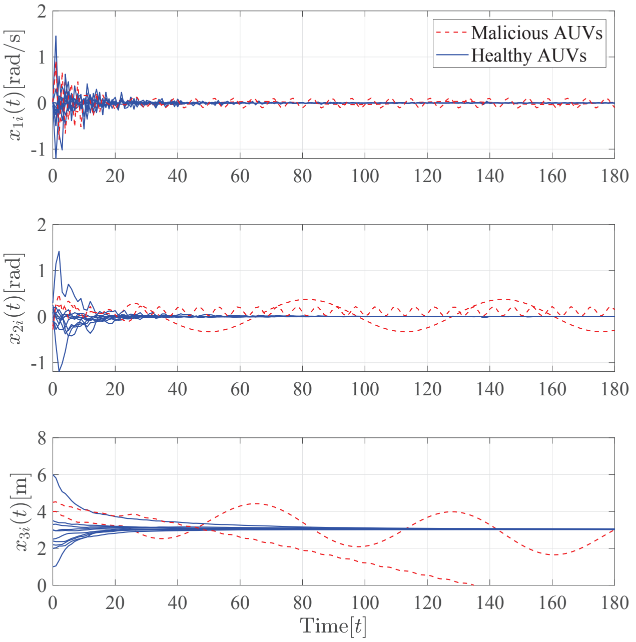

Fig. 6. State trajectories of the agents under the consensus protocol of Theorem 2 when the communication links fail stochastically and the communication channels are subject to noises.

Despite the existing studies in the literature on resilient consensus control of MASs, the communication network was considered unreliable because of stochastic links failure and noises. Accordingly, upon some conditions on the stochastic probabilities of the unreliable communication network, achieving almost sure safe consensus among the healthy agents was guaranteed.

The main assumption of the proposed strategy was model-transformation based on the knowledge of the agents models and parameters. Thus, extension of the results to networks of agents with more general models and with model and parameter uncertainties is a topic of research to be studied as future work. Moreover, the proposed consensus strategy was based on states feedback, and achieving consensus based on only outputs feedback is another open problem in this area.

Appendix

Proof: [Proof of Theorem 1] The proof is carried out in two steps:

- First, according to Part (a) of the consensus control strategy and (9), we show that the safety variables $s_i(t), i \in \mathcal{V}_h$, remain in a convex set depending on $s_i(0), i \in \mathcal{V}_h$. Then, based on the design of $s_i(t), i \in \mathcal{V}_h$, the boundedness of the agents states is guaranteed.
- In the next step, we show that if $\mathbb{P}\{\mathcal{G}(p)$ is $(2f+1)$-robust$\} \neq 0$, while $s_i(t), i \in \mathcal{V}_h$, remain in a set/bound, they almost surely converge toward a common value. Then, the achievement of almost sure consensus in the MAS as (5) is concluded.

Step 1- By substituting (9) into (3) and by considering (7), one can observe that

$$\begin{aligned} s_i(t+1) =& s_i(t) + \alpha_i(p,t) \\ &\times \sum_{j=1}^{N} a_{ij}(p,t)\big[s_j(t) - s_i(t)\big], i \in \mathcal{V}_h, \end{aligned} \tag{14}$$

which from (10), it implies that $s_i(t+1)$ is a convex combination of $s_i(t)$ and $s_j(t), j \in \mathcal{N}_i(p,t)$. To show safety in the healthy agents behaviors, at each time instant, let us define

$$\begin{aligned} s_M(t) &= \max_{i \in \mathcal{V}_h}\{s_i(t)\}, \\ s_m(t) &= \min_{i \in \mathcal{V}_h}\{s_i(t)\}. \end{aligned}$$

Since the set of the malicious agents is f-local, the healthy agent (or agents) with the safety variable $s_M(t)$ has at most f neighbors with safety variables outside the range $[s_m(t), s_M(t)]$. In this condition, according to Part (a) of the consensus control strategy, these neighbors will be ignored by the agent. Therefore, as $s_i(t+1)$ is a convex combination of $s_i(t)$ and $s_j(t), j \in \mathcal{N}_i(p,t)$, we have

$$s_m(t) \leq s_M(t+1) \leq s_M(t). \tag{15}$$

We have similar arguments for the healthy agent (or agents) with safety variable $s_m(t)$ as well, and therefore

$$s_m(t) \leq s_m(t+1) \leq s_M(t). \tag{16}$$

From (15) and (16), it follows that $s_M(t)$ is nonincreasing over time, while $s_m(t)$ is nondecreasing, and accordingly, $s_i(t)$ is always bounded as $s_m(0) \leq s_i(t) \leq s_M(0)$. Note that the healthy agent (or agents) with maximum/minimum safety variable may not be unique over time but the maximum/minimum value of $s_i(t), i \in \mathcal{V}_h$, is always nonincreasing/nondecreasing. From (3) and (6), it follows that $\xi_{m,i}(t) = \xi_{0,i}(t+m), m \in \{1, 2, \dots, n-1\}$. Thus, based on the definition of the safety variable $s_i(t)$ in (7), we have

$$\zeta_i(t+1) = \acute{A}\zeta_i(t) + \acute{B}s_i(t) \tag{17}$$

in which

$$\zeta_i(t) = \begin{bmatrix}\xi_{n-2,i}(t) & \xi_{n-3,i}(t) & \dots & \xi_{1,i}(t) & \xi_{0,i}(t)\end{bmatrix}^\top,$$

and

$$\acute{A} = \begin{bmatrix} -\lambda_1 & -\lambda_2 & \dots & -\lambda_{n-2} & -\lambda_{n-1} \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}, \acute{B} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

where due the Schur stability of the polynomial (8), $\acute{A}$ is Schur stable. In this condition, because of the input to state stability of Schur stable systems [49], the input to state stability of the system (17) with respect to the input $s_i(t)$ can be observed. Now, since $s_m(0) \leq s_i(t) \leq s_M(0)$, the boundedness of the healthy agents states can be concluded which the bound depends on the initial states.

Step 2- Now, we should show achieving almost sure consensus in the MAS. Because of the stochastic properties of the interaction links, although (15) and (16) are satisfied, the convergence of $s_M(t)$ and $s_m(t)$ are probabilistic. By defining

$$s(t) = \begin{bmatrix} s_1(t) & s_2(t) & \dots & s_N(t) \end{bmatrix}^\top,$$

and by considering the filtration

$$\mathcal{F}_r = \{s(0), s(1), \dots, s(r)\}, r \geq 0, \tag{18}$$

from (15) and (16), it can be said that

$$\begin{aligned} &\mathbb{E}\big\{|s_M(r) - s_m(r)|\big\} < \infty, \forall r, \\ &\mathbb{E}\{s_M(r) - s_m(r)|\mathcal{F}_{r-1}\} \leq s_M(r-1) - s_m(r-1), r \geq 1. \end{aligned}$$

Now, since $s_M(t) - s_m(t)$ is $\mathcal{F}_r$-measurable, it satisfies the conditions of super-martingales given in Section II-C. Therefore, from the super-martingales convergence theorem [36], it follows that the limit of $s_M(t) - s_m(t)$ almost surely exists such that

$$\lim_{t\to\infty} \big[s_M(t) - s_m(t)\big] \xrightarrow{\text{a.s.}} s_f \tag{19}$$

where $s_f \in \mathbb{R}_{\geq 0}$. If $s_f \neq 0$, at each time instant, we can consider three sets of the healthy agents defined as $\mathcal{S}_1(t)$, $\mathcal{S}_2(t)$, and $\mathcal{S}_3(t)$ such that $\mathcal{S}_1(t)$ contains all the agents with maximum $s_i(t), i \in \mathcal{V}_h$, $\mathcal{S}_2(t)$ contains all the agents with minimum $s_i(t), i \in \mathcal{V}_h$, and $\mathcal{S}_1(t) \cup \mathcal{S}_2(t) \cup \mathcal{S}_3(t) = \mathcal{V}_h$. In other words,

$$\begin{aligned} \mathcal{S}_1(t) &= \big\{i \in \mathcal{V}_h | s_i(t) = s_M(t)\big\}, \\ \mathcal{S}_2(t) &= \big\{i \in \mathcal{V}_h | s_i(t) = s_m(t)\big\}, \\ \mathcal{S}_3(t) &= \mathcal{V}_h \backslash (\mathcal{S}_1(t) \cup \mathcal{S}_2(t)). \end{aligned} \tag{20}$$

Since $\mathbb{P}\big\{\mathcal{G}(p)$ is $(2f+1)$-robust$\big\} \neq 0$, the probability of the $(2f{+}1)$-reachability of one of the sets $\mathcal{S}_1(t)$ and $\mathcal{S}_2(t)$ is nonzero. Therefore, at each time instant, with a nonzero probability there is at least one healthy agent in $\mathcal{S}_1(t)$ or $\mathcal{S}_2(t)$ which has at least $2f{+}1$ neighbors outside its set. Since the malicious agents set is f-local, with a nonzero probability this agent has at least $f{+}1$ healthy neighbors outside its set and according to (20) and based on the selection strategy in Part (a), it uses the information of at least one of them. Thus, by considering (14), with a nonzero probability this agent converges toward the other set. Note that by considering (10) and (14), $s_i(t+1)$ is a convex combination of $s_i(t)$ and $s_j(t), j \in \mathcal{N}_i(p,t)$, and according to the selection strategy in Part (a), possible malicious agents with the safety variables outside the range $[s_m(t), s_M(t)]$ will be ignored by the healthy agents. Therefore, while $s_f \neq 0$, the convex range will be shortened over time such that

$$\lim_{t\to\infty} \big[s_M(t) - s_m(t)\big] \xrightarrow{\text{i.p.}} 0. \tag{21}$$

From (21), it follows that (19) can be satisfied only if $s_f = 0$. Thus, as $s_M(t)$ is always nonincreasing and $s_m(t)$ is always nondecreasing, there exists an *a priori* unknown finite constant $s_a \in \mathbb{R}$ such that

$$\lim_{t\to\infty} s_i(t) \xrightarrow{\text{a.s.}} s_a, i \in \mathcal{V}_h. \tag{22}$$

If we define an error vector $e_i(t)$ as

$$e_i(t) = \zeta_i(t) - \frac{s_a \mathbb{1}_{n-1}}{1 + \sum_{m=1}^{n-1} \lambda_m},$$

by considering (17) and (22), one gets

$$e_i(t+1) = \acute{A}e_i(t) + \acute{B}\varepsilon_i(t) \tag{23}$$

where $\varepsilon_i(t) = s_i(t) - s_a$, and $\lim_{t\to\infty} \varepsilon_i(t) \xrightarrow{\text{a.s.}} 0$. Based on the Lyapunov stability criterion for linear systems, since $\acute{A}$ is Schur stable, for each symmetric positive definite $Q \in \mathbb{R}^{(n-1)\times(n-1)}$, there exists a symmetric positive definite $P \in \mathbb{R}^{(n-1)\times(n-1)}$ such that $\acute{A}^\top P \acute{A} - P = -Q$. Thus, for a Q, we consider a Lyapunov candidate of the error vector as $V_i(t) = e_i(t)^\top P e_i(t)$, and accordingly along (23), it can be said that

$$\begin{aligned} V_i(t+1) - V_i(t) = &- e_i(t)^\top Q e_i(t) + 2e_i(t)^\top \acute{A}^\top P \acute{B} \varepsilon_i(t) \\ &+ \varepsilon_i(t)^2 \acute{B}^\top P \acute{B}. \end{aligned}$$

From the results of Step 1 and according to definition of $\varepsilon_i(t)$ and $e_i(t)$, it follows that there exists a finite $\eta_i(t) \in \mathbb{R}_{\geq 0}$ such that $|2e_i(t)^\top \acute{A}^\top P \acute{B} \varepsilon_i(t) + \varepsilon_i(t)^2 \acute{B}^\top P \acute{B}| \leq \eta_i(t)$. Moreover, $\eta_i(t)$ almost surely converges to zero as $\varepsilon_i(t)$ almost surely converges to zero. In this condition, for a nonzero $V_i(t)$, since $-e_i(t)^\top Q e_i(t) < 0$, it can be said that for some $d_i \in \mathbb{R}_{>0}$, there exists a finite time $t_{d_i} \in \mathbb{N}$ such that [50]

$$V_i(t+1) - V_i(t) \leq -d_i \text{ a.s.}, t \geq t_{d_i}.$$

Therefore, as $t \to \infty$, $V_i(t)$ almost surely converges to zero; thus, the consensus errors almost surely converge to zero as well, implying that

$$\lim_{t\to\infty} \xi_i(t) \xrightarrow{\text{a.s.}} \frac{s_a \mathbb{1}_n}{1 + \sum_{m=1}^{n-1} \lambda_m}.$$

Since $x_i(t) = T\xi_i(t)$, one can conclude that the objective (5) is achieved which implies achieving almost sure consensus in the MAS (1), and the proof is completed. $\blacksquare$

Proof: [Proof of Theorem 2] The main reason of using the property (12b) is that as $t \to \infty$, it omits the noisy signals in the protocol. However, these noisy signals are necessary to be employed by each healthy agent such that consensus with other healthy agents in the network is achieved. Therefore, while the noisy signals are vanishing, we should simultaneously guarantee achieving almost sure consensus in the MAS. We present the proof in two steps:

- First, considering the effect of the noises on Parts (a) and (b) of the consensus control strategy, we show that under the consensus protocol (11), $\mathbb{E}\{s_i(t)\}, i \in \mathcal{V}_h$, remain in a convex set depending on $s_i(0), i \in \mathcal{V}_h$. Then, based on the design of $s_i(t), i \in \mathcal{V}_h$, the almost sure boundedness of the agents states is guaranteed.
- In the next step, from the property (12b), we conclude that the limits of $s_i(t), i \in \mathcal{V}_h$, exist almost surely. Then, we show that if $\mathbb{P}\{\mathcal{G}(p)$ is $(2f+1)$-robust$\} \neq 0$, according to the property (12c), the almost sure limits of $s_i(t), i \in \mathcal{V}_h$, are identical. Accordingly, achieving almost sure consensus in the MAS as (5) is concluded.

Step 1- By substituting the consensus protocol (11) into (3), from the definition of $s_i(t)$, one can observe that

$$\begin{aligned} s_i(t+1) = &s_i(t) + \alpha_i(p,t,\omega) \\ &\times \sum_{j=1}^{N} a_{ij}(p,t,\omega)\big[\tilde{s}_{ij}(t) - s_i(t)\big], i \in \mathcal{V}_h. \end{aligned} \tag{24}$$

Based on (24) and according to (12a), the healthy agents safety variables will be updated such that $s_i(t+1)$ will be a convex combination of $s_i(t)$ and $\tilde{s}_{ij}(t), j \in \mathcal{N}_i(p,t,\omega)$, while the set $\mathcal{N}_i(p,t,\omega)$ is determined with stochastic errors due to noises. Indeed, in stochastic analysis of the evolution of $s_i(t), i \in \mathcal{V}_h$, we should note that some $a_{ij}(p,t,\omega), j \in \mathcal{N}_i(p)$, may not be the same as them when there are no noises in the communication channels, i.e., for some $j \in \mathcal{N}_i(p)$, $a_{ij}(p,t,\omega) \neq a_{ij}(p,t)$. To analyze the evolution of $s_i(t), i \in \mathcal{V}_h$, at each time instant, we first assume a case when Part (a) of the consensus control strategy is not affected by noises, while Part (b) is affected (let us call it Case I). In other words, we assume a case when the adjacency matrix is not affected by noises, while the interaction term of (24) denoted by $\big[\tilde{s}_{ij}(t) - s_i(t)\big]$ is affected. Although this assumption is not realistic, we will use the obtained results for the real case when both parts are affected by noises (let us call it Case II). Indeed, we will show that if Case I is satisfied, the MAS has stochastic properties the same as when there are no noises in the network. Then, we will show that if Case II is satisfied, the MAS stochastic trajectory will be in a set determined by Case I. Thus, we continue with Case I. Let us consider a filtration the same as (18). Therefore, if Case I is satisfied, we have

$$\begin{aligned} &\mathbb{E}\Big\{\alpha_i(p,t,\omega)a_{ij}(p,t,\omega)\big[\tilde{s}_{ij}(t) - s_i(t)\big]\Big|\mathcal{F}_t\Big\} \\ &= \mathbb{E}\Big\{\alpha_i(p,t)a_{ij}(p,t)\big[\tilde{s}_{ij}(t) - s_i(t)\big]\Big|\mathcal{F}_t\Big\}, \end{aligned}$$

and as the communication noises are MDSs, one can observe that

$$\begin{aligned} &\mathbb{E}\Big\{\alpha_i(p,t,\omega)a_{ij}(p,t,\omega)\big[\tilde{s}_{ij}(t) - s_i(t)\big]\Big|\mathcal{F}_t\Big\} \\ &= \mathbb{E}\{\alpha_i(p,t)a_{ij}(p,t)|\mathcal{F}_t\}\big[s_j(t) - s_i(t)\big]. \end{aligned} \tag{25}$$

Therefore, from (12a) and (24), we can say that $\mathbb{E}\{s_i(t+1)|\mathcal{F}_t\}$ is a convex combination of $s_i(t)$ and $s_j(t), j \in \mathcal{N}_i(p,t)$, and thus based on similar arguments in the proof of Theorem 1 for (15) and (16),

$$\begin{aligned} s_m(t) \le \mathbb{E}\{s_M(t+1)|\mathcal{F}_t\} \le s_M(t), \\ s_m(t) \le \mathbb{E}\{s_m(t+1)|\mathcal{F}_t\} \le s_M(t). \end{aligned} \tag{26}$$

In this case, based on the selection strategy of Part (a), we have four sets of neighboring agents for the ith healthy agent as follows:

$$\begin{aligned} \mathcal{I}_{h,i}(p,t) &= \{j \in \mathcal{N}_i(p)|s_j(t) > s_i(t), a_{ij}(p,t) = 0\}, \\ \mathcal{S}_{h,i}(p,t) &= \{j \in \mathcal{N}_i(p)|s_j(t) \ge s_i(t), a_{ij}(p,t) > 0\}, \\ \mathcal{S}_{l,i}(p,t) &= \{j \in \mathcal{N}_i(p)|s_j(t) \le s_i(t), a_{ij}(p,t) > 0\}, \\ \mathcal{I}_{l,i}(p,t) &= \{j \in \mathcal{N}_i(p)|s_j(t) < s_i(t), a_{ij}(p,t) = 0\} \end{aligned} \tag{27}$$

where

$$\mathcal{I}_{l,i}(p,t) \le \mathcal{S}_{l,i}(p,t) \le \mathcal{S}_{h,i}(p,t) \le \mathcal{I}_{h,i}(p,t).$$

In the presence of communication noises, a communicated information may jump to incorrect values at some time instants. Therefore, if Case II is satisfied, considering the effects of noises on the selection strategy of Part (a) may lead to jumps of some agents among the mentioned four sets, and for some $j \in \mathcal{N}_i(p)$, we may have $a_{ij}(p,t,\omega) \neq a_{ij}(p,t)$. Thus, for Case II, due to jumps among the four sets (27), we can define the sets

$$\begin{aligned} \mathcal{I}_{h,i}(p,t,\omega) &= \{j \in \mathcal{N}_i(p)|\tilde{s}_{ij}(t) > s_i(t), a_{ij}(p,t,\omega) = 0\}, \\ \mathcal{S}_{h,i}(p,t,\omega) &= \{j \in \mathcal{N}_i(p)|\tilde{s}_{ij}(t) \ge s_i(t), a_{ij}(p,t,\omega) > 0\}, \\ \mathcal{S}_{l,i}(p,t,\omega) &= \{j \in \mathcal{N}_i(p)|\tilde{s}_{ij}(t) \le s_i(t), a_{ij}(p,t,\omega) > 0\}, \\ \mathcal{I}_{l,i}(p,t,\omega) &= \{j \in \mathcal{N}_i(p)|\tilde{s}_{ij}(t) < s_i(t), a_{ij}(p,t,\omega) = 0\} \end{aligned}$$

where

$$\mathcal{I}_{l,i}(p,t,\omega) \le \mathcal{S}_{l,i}(p,t,\omega) \le \mathcal{S}_{h,i}(p,t,\omega) \le \mathcal{I}_{h,i}(p,t,\omega).$$

Note that when Case II is satisfied, $\alpha_i(p,t,\omega)a_{ij}(p,t,\omega)$ and $\big[\tilde{s}_{ij}(t) - s_i(t)\big]$ may not be stochastically uncorrelated and thus results like (25) may not hold. To analyze the MAS behavior when Case II is satisfied, we should investigate the effects of jumps among the four sets (27) on the stochastic inequality (26). Any jump from $\mathcal{S}_{h,i}(p,t) \cup \mathcal{S}_{l,i}(p,t)$ to $\mathcal{I}_{h,i}(p,t) \cup \mathcal{I}_{l,i}(p,t)$ leads to ignoring some neighbors which were not ignored in Case I. Based on the selection strategy in Part (a), we have $|\mathcal{I}_{h,i}(p,t,\omega)| \le f$ and $|\mathcal{I}_{l,i}(p,t,\omega)| \le f$; thus, ignoring some neighbors which were selected in Case I may lead to selecting some neighbors which were ignored in Case I. In these conditions, by ignoring and selecting some new neighbors, for $\tilde{s}_{ij}(t) \ge s_i(t), j \in \mathcal{N}_i(p,t,\omega)$, compared with (25), we should have (note that compared with (25), neighbors with larger $\tilde{s}_{ij}(t)$ can be ignored)

$$\begin{aligned}0 &\le \sum_{j=1}^{N} \mathbb{E}\Big\{\alpha_i(p,t,\omega)a_{ij}(p,t,\omega)\big[\tilde{s}_{ij}(t)-s_i(t)\big]\Big|\mathcal{F}_t\Big\}\\ &\le \sum_{j=1}^{N} \mathbb{E}\Big\{\alpha_i(p,t)a_{ij}(p,t)\big[\tilde{s}_{ij}(t)-s_i(t)\big]\Big|\mathcal{F}_t\Big\}\end{aligned}$$

implying that

$$\begin{aligned}0 &\le \sum_{j=1}^{N} \mathbb{E}\Big\{\alpha_i(p,t,\omega)a_{ij}(p,t,\omega)\big[\tilde{s}_{ij}(t)-s_i(t)\big]\Big|\mathcal{F}_t\Big\}\\ &\le \sum_{j=1}^{N} \mathbb{E}\{\alpha_i(p,t)a_{ij}(p,t)|\mathcal{F}_t\}\big[s_j(t)-s_i(t)\big].\end{aligned} \tag{28}$$

In a similar way, for $\tilde{s}_{ij}(t) \le s_i(t), j \in \mathcal{N}_i(p,t,\omega)$, we have

$$\begin{aligned}&\sum_{j=1}^{N} \mathbb{E}\{\alpha_i(p,t)a_{ij}(p,t)|\mathcal{F}_t\}\big[s_j(t)-s_i(t)\big] \le\\ &\sum_{j=1}^{N} \mathbb{E}\Big\{\alpha_i(p,t,\omega)a_{ij}(p,t,\omega)\big[\tilde{s}_{ij}(t)-s_i(t)\big]\Big|\mathcal{F}_t\Big\} \le 0.\end{aligned} \tag{29}$$

It was mentioned that in Case I, (25) guarantees (26); thus, in Case II, (28) and (29) also ensure (26). Now, we investigate the effects of other types of jumps. By any jump from $\mathcal{I}_{h,i}(p,t) \cup \mathcal{I}_{l,i}(p,t)$ to $\mathcal{S}_{h,i}(p,t) \cup \mathcal{S}_{l,i}(p,t)$, some neighbors which were not selected in Case I will be selected. Based on the selection strategy in Part (a), $|\mathcal{I}_{h,i}(p,t,\omega)| < f$ only if $\{j \in \mathcal{S}_{h,i}(p,t,\omega)|\tilde{s}_{ij}(t) > s_i(t)\} = \varnothing$, and $|\mathcal{I}_{l,i}(p,t,\omega)| < f$ only if $\{j \in \mathcal{S}_{l,i}(p,t,\omega)|\tilde{s}_{ij}(t) < s_i(t)\} = \varnothing$; therefore, selecting some neighbors which were ignored in Case I may lead to ignoring some neighbors which were selected in Case I. In these conditions, by ignoring and selecting some new neighbors, for $\tilde{s}_{ij}(t) \ge s_i(t), j \in \mathcal{N}_i(p,t,\omega)$, and for $\tilde{s}_{ij}(t) \le s_i(t), j \in \mathcal{N}_i(p,t,\omega)$, compared with (25), we should still have (28) and (29), respectively, ensuring (26). The third type of jumps is from $\mathcal{I}_{h,i}(p,t) \cup \mathcal{I}_{l,i}(p,t)$ to $\mathcal{I}_{h,i}(p,t) \cup \mathcal{I}_{l,i}(p,t)$ under which compared with Case I, the members of $\mathcal{I}_{h,i}(p,t)$ and $\mathcal{I}_{l,i}(p,t)$ may change. On the one hand, $|\mathcal{I}_{h,i}(p,t,\omega)| < f$ only if $\{j \in \mathcal{S}_{h,i}(p,t,\omega)|\tilde{s}_{ij}(t) > s_i(t)\} = \varnothing$, and $|\mathcal{I}_{l,i}(p,t,\omega)| < f$ only if $\{j \in \mathcal{S}_{l,i}(p,t,\omega)|\tilde{s}_{ij}(t) < s_i(t)\} = \varnothing$; one the other hand, we always have $|\mathcal{I}_{h,i}(p,t,\omega)| \le f$ and $|\mathcal{I}_{l,i}(p,t,\omega)| \le f$. Therefore, based on the arguments the same as those for the previous two types of jumps, for $\tilde{s}_{ij}(t) \ge s_i(t), j \in \mathcal{N}_i(p,t,\omega)$, and for $\tilde{s}_{ij}(t) \le s_i(t), j \in \mathcal{N}_i(p,t,\omega)$, compared with (25), we should still have (28) and (29), respectively, ensuring (26). Finally, if we have jumps from $\mathcal{S}_{h,i}(p,t) \cup \mathcal{S}_{l,i}(p,t)$ to $\mathcal{S}_{h,i}(p,t) \cup \mathcal{S}_{l,i}(p,t)$, it implies (25), and the stochastic inequality (26) still is satisfied. By considering the combination of the mentioned types of jumps among the four sets (27), if Case II is satisfied; then, the stochastic inequality (26) is satisfied. As a result

$$s_m(0) \le \mathbb{E}\{s_i(t+1)|\mathcal{F}_t\} \le s_M(0)$$

which implies that $|s_i(t)| < \infty$ a.s., because if the boundedness is not almost surely, there exist nonzero probabilities for instability which is in contradiction with the bounded expectation of $s_i(t)$. Now, if we consider the evolution of $\zeta_i(t)$ along the Schur stable system (17), due to the input to state stability of Schur stable systems [49], one gets

$$\|\zeta_i(t)\| < \infty \text{ a.s..} \tag{30}$$

Step 2- Due to the MDS properties of the communication noises, based on arguments similar to those for $s_i(t)$, it can be said that they are almost surely bounded. In this condition, the property (12b) along (30) results in

$$\lim_{t\to\infty} \alpha_i(p,t,\omega)\sum_{j=1}^{N} a_{ij}(p,t,\omega)\big[\tilde{s}_{ij}(t)-s_i(t)\big] \xrightarrow{\text{a.s.}} 0,$$

and from (24), it follows that

$$\lim_{t\to\infty} \big[s_i(t+1)-s_i(t)\big] \xrightarrow{\text{a.s.}} 0. \tag{31}$$

Therefore, (30) and (31) imply that the limit of $s_i(t)$ exists almost surely such that

$$\lim_{t\to\infty} s_i(t) \xrightarrow{\text{a.s.}} s_{ai} \tag{32}$$

where $s_{ai} \in \mathbb{R}$. Now, we show that $s_{ai} = s_{aj}, i,j \in \mathcal{V}_h$. To achieve this goal, we show that if $s_{ai} \ne s_{aj}$ for some $i,j \in \mathcal{V}_h$, we reach a contradiction with (32). From (32), it follows that for any neighborhood $\mu_i \in \mathbb{R}_{>0}$ close to s_{ai}, there exists a finite time $t_{\mu_i} \in \mathbb{N}$ such that

$$|s_i(t) - s_{ai}| \le \mu_i \text{ a.s.}, t \ge t_{\mu_i}. \tag{33}$$

Moreover, from (32), if for some $i,j \in \mathcal{V}_h$, $s_{ai} \ne s_{aj}$, there exists a finite time $t_{\mu_i,j} \ge t_{\mu_i}$ such that

$$|s_j(t) - s_i(t)| > 0 \text{ a.s.}, t \ge t_{\mu_i,j}. \tag{34}$$

For $t \ge \max_{i,j}\{t_{\mu_i,j}\}$, let us consider three sets of the healthy agents defined as $\mathcal{S}_1$, $\mathcal{S}_2$, and $\mathcal{S}_3$ such that $\mathcal{S}_1$ contains all the healthy agents with maximum $s_{ai}, i \in \mathcal{V}_h$, $\mathcal{S}_2$ contains all the healthy agents with minimum $s_{ai}, i \in \mathcal{V}_h$, and $\mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3 = \mathcal{V}_h$. In other words,

$$\begin{aligned}\mathcal{S}_1 &= \big\{i \in \mathcal{V}_h | s_{ai} = \max_{j\in\mathcal{V}_h}\{s_{aj}\}\big\},\\ \mathcal{S}_2 &= \big\{i \in \mathcal{V}_h | s_{ai} = \min_{j\in\mathcal{V}_h}\{s_{aj}\}\big\},\\ \mathcal{S}_3 &= \mathcal{V}_h \backslash (\mathcal{S}_1 \cup \mathcal{S}_2),\end{aligned} \tag{35}$$

and as for some $i,j \in \mathcal{V}_h$, $s_{ai} \ne s_{aj}$, we have $\mathcal{S}_1 \ne \mathcal{S}_2$. Since $\mathbb{P}\big\{\mathcal{G}(p) \text{ is } (2f+1)\text{-robust}\big\} \ne 0$, the probability of the $(2f{+}1)$-reachability of one of the sets $\mathcal{S}_1$ and $\mathcal{S}_2$ is nonzero. Therefore, at each time instant, with a nonzero probability there exists at least one healthy agent in $\mathcal{S}_1$ or $\mathcal{S}_2$ which has at least $2f{+}1$ neighbors outside its set. Since the malicious agents set is f-local, with a nonzero probability this agent has at least $f{+}1$ healthy neighbors outside its set and according to (35) and based on the selection strategy in Part (a), it uses the information of at least one of these healthy neighbors. Therefore, it can be said that with a nonzero probability there exists at least one healthy agent $i \in \mathcal{S}_1$ or $i \in \mathcal{S}_2$ which uses the information of at least one healthy agent $j \in \mathcal{V}_h$ outside its set which satisfies the inequalities (34) as follows:

$$\begin{aligned}&s_j(t) - s_i(t) < 0 \text{ a.s.}, t \ge \max_{i,j}\{t_{\mu_i,j}\} \qquad i \in \mathcal{S}_1\\ &s_j(t) - s_i(t) > 0 \text{ a.s.}, t \ge \max_{i,j}\{t_{\mu_i,j}\} \qquad i \in \mathcal{S}_2.\end{aligned} \tag{36}$$

According to (12a) and (12c), for any $\phi_i(t) \in \mathbb{R}_{>0}$ and any finite constant $\chi_i \in \mathbb{R}_{>0}$, there exists a finite time period $[t_0, t_0 + \tau_i], \tau_i \in \mathbb{N}$, such that

$$\sum_{t=t_0}^{t_0+\tau_i} \alpha_i(p, t, \omega)\phi_i(t) \geq \chi_i. \tag{37}$$

On the other hand, according to (36) and considering the stochastic properties of the communication links and noises, there exists a nonzero probability such that over the period $[t_0, t_0+\tau_i]$, $t_0 \geq \max_{i,j}\{t_{\mu_i,j}\}$ (based on the selection strategy in Part (a) and since $\mathcal{V}_m$ is f-local),

$$\begin{aligned} &\sum_{j=1}^{N} a_{ij}(p,t,\omega)\big[\tilde{s}_{ij}(t) - s_i(t)\big] < 0 \qquad i \in \mathcal{S}_1 \\ &\sum_{j=1}^{N} a_{ij}(p,t,\omega)\big[\tilde{s}_{ij}(t) - s_i(t)\big] > 0 \qquad i \in \mathcal{S}_2. \end{aligned} \tag{38}$$

In this condition, considering (37) and (38) in the protocol (24), there exists a nonzero probability such that $s_i(t)$ leaves a neighborhood determined in (33) which has a contradiction with almost sure convergence of $s_i(t)$ described in (32). Therefore, we should have $s_{ai} = s_{aj}, i, j \in \mathcal{V}_h$, and therefore from (32), we have

$$\lim_{t\to\infty} s_i(t) \xrightarrow{\text{a.s.}} s_a \tag{39}$$

where $s_a \in \mathbb{R}$. By considering (30) and (39), following a procedure similar to the proof of Theorem 1 for (22), almost sure convergence of the healthy agents upon their state vectors will be concluded. ■

References

[1] Z. Dang and Y. Zhang, “Control design and analysis of an innerformation flying system,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 51, no. 3, pp. 1621–1634, Jul. 2015.

[2] H. Rezaee and F. Abdollahi, “Attitude consensusability in multispacecraft systems using magnetic actuators,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 1, pp. 513–519, Feb. 2017.

[3] W. Ren, “Consensus strategies for cooperative control of vehicle formations,” *IET Control Theory Appl.*, vol. 1, no. 2, pp. 505–512, Mar. 2007.

[4] H. Rezaee and F. Abdollahi, “Pursuit formation of double-integrator dynamics using consensus control approach,” *IEEE Trans. Ind. Electron.*, vol. 62, no. 7, pp. 4249–4256, Jul. 2015.

[5] B. Chen, G. Pin, W. M. Ng, T. Parisini, and S. R. Hui, “A fast-convergent modulation integral observer for online detection of the fundamental and harmonics in grid-connected power electronics systems,” *IEEE Trans. Power Electron.*, vol. 32, no. 4, pp. 2596–2607, Apr. 2017.

[6] M. U. Qureshi and S. Grijalva, “Enhanced frequency response based on multiagent distributed power agreement,” *IEEE Trans. Ind. Appl.*, vol. 54, no. 2, pp. 1746–1755, Mar./Apr. 2018.

[7] R. Olfati-Saber and R. M. Murray, “Consensus problems in networks of agents with switching topology and time-delays,” *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1520–1533, Sep. 2004.

[8] Y. Hatano and M. Mesbahi, “Agreement over random networks,” *IEEE Trans. Autom. Control*, vol. 50, no. 11, pp. 1867–1872, Nov. 2005.

[9] W. Ren, “On consensus algorithms for double-integrator dynamics,” *IEEE Trans. Autom. Control*, vol. 53, no. 6, pp. 1503–1509, Jul. 2008.

[10] H. Rezaee and F. Abdollahi, “Average consensus over high-order multiagent systems,” *IEEE Trans. Autom. Control*, vol. 60, no. 11, pp. 3047–3052, Nov. 2015.

[11] E. Franco, R. Olfati-Saber, T. Parisini, and M. M. Polycarpou, “Distributed fault diagnosis using sensor networks and consensus-based filters,” in *Proc. 45th IEEE Conf. Decis. Control*, San Diego, CA, USA, Dec. 2006, pp. 386–391.

[12] F. Boem, A. J. Gallo, G. Ferrari-Trecate, and T. Parisini, “A distributed attack detection method for multi-agent systems governed by consensus-based control,” in *Proc. 56th IEEE Conf. Decis. Control*, Melbourne, VIC, Australia, Dec. 2017, pp. 5961–5966.

[13] A. D’Innocenzo, F. Smarra, and M. D. Di Benedetto, “Resilient stabilization of multi-hop control networks subject to malicious attacks,” *Automatica*, vol. 71, pp. 1–9, Sep. 2016.

[14] A. Lu and G. Yang, “Input-to-state stabilizing control for cyber-physical systems with multiple transmission channels under denial-of-service,” *IEEE Trans. Autom. Control*, vol. 63, no. 6, pp. 1813–1820, Jun. 2018.

[15] F. Pasqualetti, A. Bicchi, and F. Bullo, “On the security of linear consensus networks,” in *Proc. 48th IEEE Conf. Decis. Control held jointly with 28th Chinese Control Conf.*, Shanghai, China, Dec. 2009, pp. 4894–4901.

[16] ——, “Consensus computation in unreliable networks: A system theoretic approach,” *IEEE Trans. Autom. Control*, vol. 57, no. 1, pp. 90–104, Jan. 2012.

[17] H. Zhang and S. Sundaram, “Robustness of information diffusion algorithms to locally bounded adversaries,” in *Proc. Amer. Control Conf.*, Montreal, QC, Canada, Jun. 2012, pp. 5855–5861.

[18] 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, Apr. 2013.

[19] S. M. Dibaji and H. Ishii, “Resilient consensus of double-integrator multi-agent systems,” in *Proc. Amer. Control Conf.*, Portland, OR, USA, Jun. 2014, pp. 5139–5144.

[20] S. M. Dibaji, H. Ishii, and R. Tempo, “Resilient randomized quantized consensus,” *IEEE Trans. Autom. Control*, vol. 63, no. 8, pp. 2508–2522, Aug. 2018.

[21] Y. Wu and X. He, “Secure consensus control for multiagent systems with attacks and communication delays,” *IEEE/CAA J. Autom. Sin.*, vol. 4, no. 1, pp. 136–142, Jan. 2017.

[22] H. J. LeBlanc and X. Koutsoukos, “Resilient first-order consensus and weakly stable, higher order synchronization of continuous-time networked multi-agent systems,” *IEEE Trans. Control Netw. Syst.*, vol. 5, no. 3, pp. 1219–1231, Sep. 2018.

[23] S. M. Dibaji and H. Ishii, “Resilient consensus of second-order agent networks: Asynchronous update rules with delays,” *Automatica*, vol. 81, pp. 123–132, Jul. 2017.

[24] D. Senejohnny, S. Sundaram, C. De Persis, and P. Tesi, “Resilience against misbehaving nodes in self-triggered coordination networks,” in *Proc. 57th IEEE Conf. Decis. Control*, Miami Beach, FL, USA, Dec. 2018, pp. 2848–2853.

[25] M. Nakamura, H. Ishii, and S. M. Dibaji, “Maximum-based consensus and its resiliency,” in *Proc. 7th IFAC Workshop Distrib. Estimation Control Netw. Syst.*, Groningen, The Netherlands, Aug. 2018, pp. 283–288.

[26] J. Usevitch and D. Panagou, “Resilient leader-follower consensus to arbitrary reference values,” in *Proc. Amer. Control Conf.*, Milwaukee, WI, USA, Jun. 2018, pp. 1292–1298.

[27] W. Abbas, Y. Vorobeychik, and X. Koutsoukos, “Resilient consensus protocol in the presence of trusted nodes,” in *Proc. 7th Int. Symp. Resilient Control Syst.*, Denver, CO, USA, Aug. 2014, pp. 1–7.

[28] J. He, P. Cheng, L. Shi, and J. Chen, “SATS: Secure average-consensus-based time synchronization in wireless sensor networks,” *IEEE Trans. Signal Process.*, vol. 61, no. 24, pp. 6387–6400, Dec. 2013.

[29] W. Zeng and M. Chow, “Resilient distributed control in the presence of misbehaving agents in networked control systems,” *IEEE Trans. Cybern.*, vol. 44, no. 11, pp. 2038–2049, Nov. 2014.

[30] R. Gowaikar, B. Hochwald, and B. Hassibi, “Communication over a wireless network with random connections,” *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 2857–2871, Jul. 2006.

[31] M. Khalili, X. Zhang, M. M. Polycarpou, T. Parisini, and Y. Cao, “Distributed adaptive fault-tolerant control of uncertain multi-agent systems,” in *Proc. 9th IFAC Symp. Fault Detection, Supervision Safety Technical Processes*, Paris, France, Sep. 2015, pp. 66–71.

[32] M. Khalili, X. Zhang, Y. Cao, M. M. Polycarpou, and T. Parisini, “Distributed adaptive fault-tolerant leader-following formation control of nonlinear uncertain second-order multi-agent systems,” *Int. J. Robust Nonlinear Control*, vol. 28, no. 15, pp. 4287–4308, Oct. 2018.

[33] Y. Wang, Y. Song, M. Krstic, and C. Wen, “Fault-tolerant finite time consensus for multiple uncertain nonlinear mechanical systems under single-way directed communication interactions and actuation failures,” *Automatica*, vol. 63, pp. 374–383, Jan. 2016.

[34] D. Ye, M. Chen, and H. Yang, “Distributed adaptive event-triggered fault-tolerant consensus of multiagent systems with general linear dynamics,” *IEEE Trans. Cybern.*, vol. 49, no. 3, pp. 757–767, Mar. 2019.

[35] C. Shi, G. Yang, and X. Li, "Data-based fault-tolerant consensus control for uncertain multiagent systems via weighted edge dynamics," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 49, no. 12, pp. 2548–2558, Dec. 2019.

[36] D. Williams, *Probability With Martingales*. Cambridge, UK: Cambridge University Press, 1991.

[37] L. Ouchti, "On the rate of convergence in the central limit theorem for martingale difference sequences," *Annales de l'Institut Henri Poincare (B) Probability and Statistics*, vol. 41, no. 1, pp. 35–43, Jan./Feb. 2005.

[38] R. Tempo, G. Calafiore, and F. Dabbene, *Randomized Algorithms for Analysis and Control of Uncertain Systems: With Applications*, 2nd ed. London, UK: Springer-Verlag London, 2013.

[39] F. Fagnani and S. Zampieri, "Randomized consensus algorithms over large scale networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 4, pp. 634–649, May 2008.

[40] A. Tahbaz-Salehi and A. Jadbabaie, "Consensus over ergodic stationary graph processes," *IEEE Trans. Autom. Control*, vol. 55, no. 1, pp. 225–230, Jan. 2010.

[41] H. Rezaee and F. Abdollahi, "Discrete-time consensus strategy for a class of high-order linear multiagent systems under stochastic communication topologies," *J. Frankl. Inst.*, vol. 354, no. 9, pp. 3690–3705, Jun. 2017.

[42] D. Vengertsev, H. Kim, J. H. Seo, and H. Shim, "Consensus of output-coupled high-order linear multi-agent systems under deterministic and Markovian switching networks," *Int. J. Syst. Sci.*, vol. 46, no. 10, pp. 1790–1799, 2015.

[43] Q. Zhang and J. Zhang, "Distributed parameter estimation over unreliable networks with Markovian switching topologies," *IEEE Trans. Autom. Control*, vol. 57, no. 10, pp. 2545–2560, Oct. 2012.

[44] L. Cheng, Y. Wang, Z. Hou, and M. Tan, "Stochastic consensus of linear multi-agent systems: Communication noises and Markovian switching topologies," in *Proc. 26th Chinese Control Decis. Conf.*, Changsha, China, May 2014, pp. 274–279.

[45] L. Wang, L. Wang, and Z. Kong, "Two controllable canonical forms for single input complex network," in *Proc. 29th Chinese Control Decis. Conf.*, Chongqing, China, May 2017, pp. 1467–1472.

[46] A. J. Gallo, M. S. Turan, F. Boem, G. Ferrari-Trecate, and T. Parisini, "Distributed watermarking for secure control of microgrids under replay attacks," in *Proc. 7th IFAC Workshop Distrib. Estimation Control Netw. Syst.*, Groningen, The Netherlands, Aug. 2018, pp. 182–187.

[47] H. Yan, J. Wang, H. Zhang, H. Shen, and X. Zhan, "Event-based security control for stochastic networked systems subject to attacks," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 50, no. 11, pp. 4643–4654, Nov. 2020.

[48] I. Saboori and K. Khorasani, "$\mathcal{H}_\infty$ consensus achievement of multi-agent systems with directed and switching topology networks," *IEEE Trans. Autom. Control*, vol. 59, no. 11, pp. 3104–3109, Nov. 2014.

[49] Z. Jiang and Y. Wang, "Input-to-state stability for discrete-time nonlinear systems," *Automatica*, vol. 37, no. 6, pp. 857–869, Jun. 2001.

[50] I. Barkana, "Defending the beauty of the Invariance Principle," *Int. J. Control*, vol. 87, no. 1, pp. 186–206, 2014.

Hamed Rezaee (S'10, M'18) received the B.Sc., M.Sc., and Ph.D. degrees in control engineering from the Department of Electrical Engineering at Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran, in 2009, 2011, and 2016, respectively. He is currently a Research Associate in the Department of Electrical and Electronic Engineering at Imperial College London, London, UK, with a research focus on resilient control and monitoring in cyber-physical systems, multiagent systems, and consensus problems.

Thomas Parisini (F'11) received the Ph.D. degree in Electronic Engineering and Computer Science in 1993 from the University of Genoa. He was with Politecnico di Milano and since 2010 he holds the Chair of Industrial Control and is Director of Research at Imperial College London. He is a Deputy Director of the KIOS Research and Innovation Centre of Excellence, University of Cyprus. Since 2001 he is also Danieli Endowed Chair of Automation Engineering with University of Trieste. In 2009-2012 he was Deputy Rector of University of Trieste. In 2018 he received an *Honorary Doctorate* from University of Aalborg, Denmark. He authored or co-authored more than 320 research papers in archival journals, book chapters, and international conference proceedings. His research interests include neural-network approximations for optimal control problems, distributed methods for cyber-attack detection and cyber-secure control of large-scale systems, fault diagnosis for nonlinear and distributed systems, nonlinear model predictive control systems and nonlinear estimation. He is a co-recipient of the IFAC Best Application Paper Prize of the Journal of Process Control, Elsevier, for the three-year period 2011-2013 and of the 2004 Outstanding Paper Award of the IEEE Trans. on Neural Networks. In 2016, he was awarded as Principal Investigator at Imperial of the H2020 European Union flagship Teaming Project KIOS Research and Innovation Centre of Excellence led by University of Cyprus. In 2012, he was awarded an ABB Research Grant dealing with energy-autonomous sensor networks for self-monitoring industrial environments. Thomas Parisini currently serves as 2020 President-Elect of the IEEE Control Systems Society and will serve as 2021-2022 President. During 2009-2016 he was the Editor-in-Chief of the IEEE Trans. on Control Systems Technology. Since 2017, he is Editor for Control Applications of Automatica and since 2018 he is the Editor in Chief of the European Journal of Control. Among other activities, he was the Program Chair of the 2008 IEEE Conference on Decision and Control and General Co-Chair of the 2013 IEEE Conference on Decision and Control. Prof. Parisini is a Fellow of the IEEE and of the IFAC.

Marios M. Polycarpou (F'06) is a Professor of Electrical and Computer Engineering and the Director of the KIOS Research Center for Intelligent Systems and Networks at the University of Cyprus. He received undergraduate degrees in Computer Science and in Electrical Engineering, both from Rice University, USA in 1987, and the M.S. and Ph.D. degrees in Electrical Engineering from the University of Southern California, in 1989 and 1992 respectively. His teaching and research interests are in intelligent systems and networks, adaptive and cooperative control systems, computational intelligence, fault diagnosis and distributed agents. Dr. Polycarpou has published more than 350 articles in refereed journals, edited books and refereed conference proceedings, and co-authored 7 books. He is also the holder of 6 patents.

Prof. Polycarpou is the recipient of the 2016 IEEE Neural Networks Pioneer Award. He is a Fellow of IEEE and IFAC, and he received with his co-authors the 2014 Best Paper Award for the journal Building and Environment (Elsevier). Prof Polycarpou served as the President of the IEEE Computational Intelligence Society (2012-2013), as the Editor-in-Chief of the IEEE Transactions on Neural Networks and Learning Systems (2004-2010), and as the President of the European Control Association (EUCA) between 2018-2019. He has participated in more than 65 research projects/grants, funded by several agencies and industry in Europe and the United States, including the prestigious European Research Council (ERC) Advanced Grant and the EU Teaming project. Prof. Polycarpou is a founding member of the Cyprus Academy of Sciences, Letters, and Arts.