

A Robustness Analysis to Structured Channel Tampering over Secure-by-design Consensus Networks

Marco Fabris and Daniel Zelazo *Senior Member, IEEE*

Abstract—This work addresses multi-agent consensus networks where adverse attackers affect the convergence performances of the protocol by manipulating the edge weights. We generalize [1] and provide guarantees on the agents’ agreement in the presence of attacks on multiple links in the network. A stability analysis is conducted to show the robustness to channel tampering in the scenario where part of the codeword, corresponding to the value of the edge weights, is corrupted. Exploiting the built-in objective coding, we show how to compensate the conservatism that may emerge because of multiple threats in exchange for higher encryption capabilities. Numerical examples related to semi-autonomous networks are provided.

Index Terms—Agents-based systems, Network analysis and control, Secure consensus protocols

I. INTRODUCTION

The consensus problem, consisting in the design of networked control algorithms under which all individuals of a given multi-agent system (MAS) attain an agreement on a certain quantity of interest [2], is commonly tackled when it is required to achieve a global prefixed task. However, due to the openness of communication protocols and the complexity of networks, the agreement of MASs may be vulnerable to malicious cyber-attacks [3]. In particular, if the agent sensors are threatened by an attacker, the measured data may be unreliable or faulty. Indeed, the attack signals can even disrupt the control performance of the group of agents through the communication topology. Therefore, resilient solutions are required to ensure that MASs fulfill consensus under security hazards [4]. Consequently, the secure control of MASs is now a crucial issue to be investigated [5]–[7].

Several recent studies illustrate the importance of giving guarantees against cyber-threats while these are attempting to disrupt a MAS that is trying to reach consensus. In [8], [9], deception attackers injecting false data are assumed to attack the agents or communication channels. To counteract this kind of disruption, classic observers, impulsive control methods and event-triggered adaptive cognitive control have been leveraged. Denial of service (DoS) attacks then represent another challenging class of cyber-threats: robust control techniques have been developed in [10]–[12] to ensure sufficient levels of agreement over MASs under DoS providing guarantees based on the maximum “quality of service” or Lyapunov theory.

M. Fabris is with the Department of Information Engineering, University of Padua, Padua, Italy. D. Zelazo is with the Faculty of Aerospace Engineering, Technion, Haifa, Israel. This research was made possible thanks to the support of Ms. Beverly Bavit, Mr. David Dibner and another anonymous donor. Corresponding author: D. Zelazo, dzelazo@technion.ac.il

As already introduced in [1], we embrace a different perspective. Instead of developing tools to secure existing networks (see [13]), we provide inherently secure embedded measures through the adoption of a network manager to guarantee robust consensus convergence. For privacy and safety concerns, such a manager is not allowed to access local states. Rather, it only intervenes in an initial phase to ensure desired convergence performance via edge weight assignment (see e.g. [14]) in a secure way. Nonetheless, differently from our previous work, this paper is meant to generalize the *secure-by-design consensus* (SBDC) dynamics towards multi-agent consensus networks where adverse attackers affect the convergence performance through a *structural* hit to the communication between manager and agents, thus corrupting edge weights in *more than a single link* of the system. We summarize our main contributions below.

- Two new guarantees based on the small gain theorem for the robust stability of the agents’ agreement are given in both continuous and discrete time domains when multiple network edges undergo a structured weight deviation from their nominal values.
- We introduce the notion of a resilience gap used to characterize the conservatism of the robustness analysis. We show that for spanning trees, the resilience gap is always zero even in multi-attack scenarios. We also discuss how this gap can be reduced by modulating the encryption capabilities used in the network.

The organization of the paper follows. Sec. II introduces the preliminary notions for multi-agent consensus and the SBDC networks. Sec. III provides a robustness analysis for the latter when subject to channel tampering modeled as multiple edge weight perturbations. A numerical case study on semi-autonomous networks is reported in Sec. IV to assess such theoretical results and, lastly, concluding remarks and future works are sketched in Sec. V.

Notation: The set of real numbers, the l -dimensional (column) vector whose elements are all ones and the l -dimensional (column) null vector are denoted by \mathbb{R} , $\mathbf{1}_l \in \mathbb{R}^l$ and $\mathbf{0}_l \in \mathbb{R}^l$, respectively, while $I_l \in \mathbb{R}^{l \times l}$ refers to as the identity matrix. Let $\Omega \in \mathbb{R}^{l \times l}$ be a square matrix. Relation $\Omega \geq 0$ means that Ω is symmetric positive semidefinite. Notation $[\Omega]_{ij}$ addresses the entry of matrix Ω at row i and column j , while Ω^\top , $\text{tr}(\Omega)$ and $\|\Omega\|$ denote respectively its transpose, its trace and its spectral norm. Operator $\text{col}_l[\Omega]$ represents the l -th column of Ω and its i -th eigenvalue is denoted by λ_i^Ω . The vector space spanned by a vector $\omega \in \mathbb{R}^l$ is denoted with $\langle \omega \rangle$. The infinity norm of ω is

identified by $\|\omega\|_\infty$. Also, $\omega = \text{vec}_{i=1}^l(\omega_i)$ defines the vectorization operator that stacks vectors ω_i , $i = 1, \dots, l$ as $\omega = [\omega_1^\top \dots \omega_l^\top]^\top$; whereas, $\text{diag}_{i=1}^l(\varsigma_i)$ is a block diagonal matrix made up with $\varsigma_i \in \mathbb{R}$, $i = 1, \dots, l$, on the diagonal and $\text{diag}(\omega) = \text{diag}_{i=1}^l(\omega_i)$. Lastly, the Kronecker product is denoted with \otimes .

II. PRELIMINARY RESULTS

Consensus models for MASs and preliminary notions are here given along with a brief overview of robustness results in consensus networks having multiple uncertain edges. The SBDC protocol is then briefly recalled.

A. Overview on uncertain consensus networks

An n -agent network can be modeled through a weighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{W})$ so that each element in the *vertex set* $\mathcal{V} = \{1, \dots, n\}$ is associated to an agent in the group, while the *edge set* $\mathcal{E} = \{e_k\}_{k=1}^m \subseteq \mathcal{V} \times \mathcal{V}$ characterizes the agents' information exchange. In addition, $\mathcal{W} = \{w_k\}_{k=1}^m$, with $m = |\mathcal{E}|$, denotes the set of weights attributed to each link. Throughout this work, bidirectional interactions among agents are assumed; hence, \mathcal{G} is an *undirected* graph. A *cycle* is defined as a non-empty and distinct sequence of edges joining a sequence of vertices, in which only the first and last vertices are equal. If a graph does not contain cycles it is said *acyclic* and if it is also connected it is called a *tree*. The *neighborhood* of node i is defined as the set $\mathcal{N}_i = \{j \in \mathcal{V} \setminus \{i\} \mid (i, j) \in \mathcal{E}\}$, while the degree of node i is defined through the cardinality $d_i = |\mathcal{N}_i|$ of neighborhood \mathcal{N}_i . Moreover, the *incidence matrix* is denoted as $E \in \mathbb{R}^{n \times m}$, in which each column $k \in \{1, \dots, m\}$ is defined via the k -th (ordered) edge $(i, j) \in \mathcal{E}$, where $i < j$ is chosen w.l.o.g., and for edge k corresponding to (i, j) one has $[E]_{lk} = -1$, if $l = i$; $[E]_{lk} = 1$, if $l = j$; $[E]_{lk} = 0$, otherwise. For all $k = 1, \dots, m$, the weight $w_k := w_{ij} = w_{ji} \in \mathbb{R}$ is related to k -th edge (i, j) , and $W = \text{diag}_{k=1}^m(w_k)$ is the diagonal matrix of edge weights. Additionally, the *Laplacian matrix*, incorporating the topological information about \mathcal{G} , is denoted as $L(\mathcal{G}) = EWE^\top$ (see [15]). Henceforward, we also suppose that graph \mathcal{G} is *connected* and $L(\mathcal{G}) \geq 0$, thus having eigenvalues λ_i^L , for $i = 1, \dots, n$, such that $0 = \lambda_1^L < \lambda_2^L \leq \dots \leq \lambda_n^L$. A sufficient condition to satisfy the latter requisite, which is adopted throughout this paper, is to take $w_{ij} > 0$ for all (i, j) . With an appropriate labeling of the edges, we can always assume that the incidence matrix $E = [E_{\mathcal{T}} \ E_{\mathcal{C}}]$ can be partitioned into the incidence matrix $E_{\mathcal{T}}$, relative to a spanning tree $\mathcal{T} \subseteq \mathcal{G}$ with $\tau = n - 1$ edges, and the incidence matrix $E_{\mathcal{C}}$, associated to $\mathcal{C} = \mathcal{G} \setminus \mathcal{T}$. Consequently, we define the cut-set matrix of \mathcal{G} (see [16]) as $R_{(\mathcal{T}, \mathcal{C})} = [I_\tau \ T_{(\mathcal{T}, \mathcal{C})}]$, with $T_{(\mathcal{T}, \mathcal{C})} = (E_{\mathcal{T}}^\top E_{\mathcal{T}})^{-1} E_{\mathcal{T}}^\top E_{\mathcal{C}}$.

A summary of the weighted consensus problem in MASs follows. Let us consider a group of n homogeneous agents, modeled by a weighted and connected graph \mathcal{G} , and assign a continuous-time state $x_i = x_i(t) \in \mathbb{R}^D$ to the i -th agent, for $i = 1, \dots, n$. The full network state is given by $\mathbf{x} = \text{vec}_{i=1}^n(x_i) \in X \subseteq \mathbb{R}^N$, with $N = nD$. The weighted consensus for a MAS can be characterized as follows.

Definition 2.1 (Weighted Consensus [15]): An n -agent network achieves *consensus* if $\lim_{t \rightarrow +\infty} \mathbf{x}(t) \in \mathcal{A}$, where, for some $\omega \in \mathbb{R}^D$, $\mathcal{A} = \langle \mathbf{1}_n \rangle \otimes \omega$ is termed the *agreement set*.

For a connected graph \mathcal{G} with positive weights, it is well known that the *linear weighted consensus protocol*, given by

$$\dot{\mathbf{x}} = -\mathbf{L}(\mathcal{G})\mathbf{x}, \quad (1)$$

where $\mathbf{L}(\mathcal{G}) = (L(\mathcal{G}) \otimes I_D)$, satisfies $\mathbf{x}(t) \in \mathcal{A}$ as $t \rightarrow +\infty$.

In this direction, we also revisit a robustness result for the consensus protocol with small-magnitude perturbations on the edge weights [16]. Within this framework, we take into consideration the perturbed Laplacian matrix $L(\mathcal{G}_{\Delta^W}) = E(W + \Delta^W)E^\top$ for a structured norm-bounded perturbation

$$\Delta^W \in \mathbf{\Delta}^W = \{\Delta^W : \Delta^W = \text{diag}_{k=1}^m(\delta_k^w), \|\Delta^W\| \leq \bar{\delta}^W\}. \quad (2)$$

Letting $\mathcal{E}_\Delta := \{e_1^\Delta, e_2^\Delta, \dots\} \subseteq \mathcal{E}$ be the (nonempty) subset of uncertain edges, we can define the matrix $P \in \{0, 1\}^{|\mathcal{E}| \times |\mathcal{E}_\Delta|}$ that selects the uncertain edges in \mathcal{E} , with $[P]_{ij} = 1$, if i and j satisfy $e_i = e_j^\Delta$; and $[P]_{ij} = 0$, otherwise. This leads to

Lemma 2.1: Let $\Delta^W \in \mathbb{R}^{|\mathcal{E}_\Delta| \times |\mathcal{E}_\Delta|}$, with Δ^W diagonal, and consider the nominal weighted consensus protocol (1). Then the perturbed consensus protocol

$$\dot{\mathbf{x}} = -(L(\mathcal{G}_{\Delta^W}) \otimes I_D)\mathbf{x} \quad (3)$$

achieves consensus $\forall \Delta^W \in \mathbf{\Delta}^W$ (defined in (2)), if

$$\|\Delta^W\| < \underbrace{\left\| P^\top R_{(\mathcal{T}, \mathcal{C})}^\top (R_{(\mathcal{T}, \mathcal{C})} W R_{(\mathcal{T}, \mathcal{C})}^\top)^{-1} R_{(\mathcal{T}, \mathcal{C})} P \right\|^{-1}}_{=: \mathcal{R}_{\mathcal{E}_\Delta}(\mathcal{G})}. \quad (4)$$

Proof: By [16, Theorem V.2], the edge-agreement version of (3) is asymptotically stable. Consequently, as \mathcal{G}_{Δ^W} is connected, one has $L(\mathcal{G}_{\Delta^W}) \geq 0$ with a simple eigenvalue at 0. Therefore, Def. 2.1 is satisfied $\forall \Delta^W \in \mathbf{\Delta}^W$. ■

In a slight abuse of convention, we refer to Lemma 2.1 as *robust stability* result; see [16] for more discussion on this notion. When $|\mathcal{E}_\Delta| = 1$, it was shown in [16] that the bound in (4) is tight and that $\mathcal{R}_{\{(u, v)\}}(\mathcal{G})$ can be interpreted as the *effective resistance* between a pair of nodes (u, v) . However, for multi-edge attacks this bound is inherently conservative; this aspect will be elaborated upon in Section III-C.

B. Secure-by-design consensus dynamics

In this work, we consider MASs that are subject to the same key principles assumed in [1], that is with presence of tasks, objective coding, information localization and a network manager¹. These elements, which comprise the basic setup of the SBDC dynamics, are recalled in the following lines.

Tasks are described by an encoded parameter θ that we term the *codeword* and the space of all tasks is denoted as Θ . Each agent in the network then decodes this objective using its *objective decoding function*, defined as $p_i : \Theta \rightarrow$

¹The network manager does not govern the agents' dynamics, namely it should not be intended to fulfill the role of a global controller. Instead, it is chosen to precisely serve as an encryption mechanism to set up and secure distributed algorithms running on the underlying MAS.

Π_i , where Π_i depends on the particular application (e.g. $\Pi_i \subseteq \mathbb{R}^n$ within the consensus setting). For $\theta \in \Theta$, $p_i(\theta)$ is called the *localized objective*. Instead, if $\theta \notin \Theta$, $p_i(\theta)$ may not be computable; nonetheless, any agent collecting such a codeword may send an alert. More precisely, the objective coding is established via the non-constant functions $p_i(\theta) : \Theta \rightarrow \Pi_i \subseteq \mathbb{R}^n$, such that $[p_i(\theta)]_j := p_{ij}(\theta)$ with

$$p_{ij}(\theta) = \begin{cases} w_{ij}, & \text{if } (i, j) \in \mathcal{E}; \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

The values w_{ij} in (5) coincide with the nominal desired consensus weights that are assigned, encoded and broadcast by the network manager, to achieve desired convergence rates or other performance metrics in consensus networks. Moreover, the information localization about the global state \mathbf{x} is expressed by means of $h_i(\mathbf{x}) : X \rightarrow Y_i \subseteq \mathbb{R}^{D \times n}$, such that $\text{col}_j[h_i(\mathbf{x})] := h_{ij}(\mathbf{x}(t)) \in \mathbb{R}^D$ with $h_{ij}(\mathbf{x}) = x_i - x_j$, if $(i, j) \in \mathcal{E}$; $h_{ij}(\mathbf{x}) = \mathbf{0}_D$, otherwise. So, the i -th agent's dynamics for the SBDC is determined by

$$\dot{x}_i = -\sum_{j \in \mathcal{N}_i} p_{ij}(\theta) h_{ij}(\mathbf{x}). \quad (6)$$

Remarkably, (6) reproduces exactly the linear consensus protocol already introduced in (1) and, defining $\mathbf{p}(\theta) = \text{vec}_{i=1}^n(p_i(\theta)) \in \mathbb{R}^{n^2}$ and $\mathbf{H}(\mathbf{x}) = \text{diag}_{i=1}^n(h_i(\mathbf{x}(t))) \in \mathbb{R}^{N \times n^2}$, dynamics (6) can be rewritten as

$$\dot{\mathbf{x}} = -\mathbf{H}(\mathbf{x})\mathbf{p}(\theta), \quad (7)$$

leading to the following theoretical result.

Lemma 2.2 ([1, Lemma III.1]): The SBDC protocol (7) reaches agreement for any given objective decoding function \mathbf{p} that satisfies (5).

Likewise, we consider and investigate the well-known discrete-time consensus dynamics described by

$$\mathbf{x}(t+1) = \mathbf{x}(t) - \epsilon \mathbf{L}(\mathcal{G})\mathbf{x}(t), \quad (8)$$

where ϵ is a common parameter shared among all agents and designed to belong to the interval $(0, 2/\lambda_n^L)$, as shown in [17]. According to the characterization in (7), the SBDC dynamics adopted in discrete time in (8) is also given by

$$\mathbf{x}(t+1) = \mathbf{x}(t) - \epsilon \mathbf{H}(\mathbf{x}(t))\mathbf{p}(\theta), \quad (9)$$

since $\mathbf{H}(\mathbf{x})\mathbf{p}(\theta) = \mathbf{L}(\mathcal{G})\mathbf{x}$ holds thanks to Lem. 2.2.

III. ROBUSTNESS OF THE SBDC PROTOCOL TO STRUCTURED CHANNEL TAMPERING

The original contribution provided by this study aims at the design of secure network systems to structured channel tampering while achieving consensus task. To this aim, the system is embedded with security measures that allow to render the network robust to small signal perturbations on some of the links. Also, a description for the structured channel tampering is given along with the relative robustness analysis for the SBDC protocol under multiple threats.

A. Models, problem statement and key assumptions

The structured channel tampering problem under analysis is formulated as follows. Similarly to the model adopted in [1], the prescribed codeword θ is subject to a perturbation $\delta^\theta \in \Delta^\theta = \{\delta^\theta : \|\delta^\theta\|_\infty \leq \rho_\Delta^\theta\}$. We let Θ be a Euclidean subspace, in particular $\Theta = \Theta_{11} \times \Theta_{12} \times \dots \times \Theta_{nn} \subseteq \mathbb{R}^{n^2}$, and allow a codeword $\theta = \text{vec}_{i=1}^n(\theta_i) \in \Theta$ to be divided into (at most) $n(n-1)/2$ relevant ‘‘subcodewords’’ $\theta^{(k)} := [\theta_i]_j = \theta_{ij}$, with $k = 1, \dots, m$, such that $\theta_{ij} = \theta_{ji}$, if $i \neq j$, and θ_{ii} is free to vary, for $i = 1, \dots, n$. Each $\theta_{ij} \in \Theta_{ij} \subseteq \mathbb{R}$ can be seen as the j -th component of the i -th codeword fragment θ_i , with $i = 1, \dots, n$. Such subcodewords influence the value of $p_{ij}(\theta)$ directly if and only if $j \in \mathcal{N}_i$, i.e., it holds that $p_{ij}(\theta) = p_{ij}(\theta_{ij})$ for all $(i, j) \in \mathcal{E}$.

Therefore, the consensus description to support this investigation is such that the i -th nominal dynamics in (7) is modified as

$$\dot{x}_i = -\sum_{j \in \mathcal{N}_i} p_{ij}(\theta_{ij} + \delta_{ij}^\theta) h_{ij}(\mathbf{x}), \quad i = 1, \dots, n, \quad (10)$$

with $\delta_{ij}^\theta = [\delta_i^\theta]_j$ and δ_i^θ satisfying $\delta^\theta = \text{vec}_{i=1}^n(\delta_i^\theta)$. Analogously, the i -th perturbed discrete time dynamics in (9) can be written as:

$$x_i(t+1) = x_i(t) - \epsilon \sum_{j \in \mathcal{N}_i} p_{ij}(\theta_{ij} + \delta_{ij}^\theta) h_{ij}(\mathbf{x}(t)), \quad (11)$$

where $\epsilon > 0$ must be selected. In light of the previous discussion, the following design problem can be now stated.

Problem 3.1: Design objective functions p_{ij} such that (10) (resp., (11) in discrete time) reaches consensus, independently from the codeword $\theta \in \Theta \subseteq \mathbb{R}^{n^2}$, while the underlying MAS is subject to a structured attack $\delta^\theta \in \Delta^\theta$ on multiple edges (belonging to \mathcal{E}_Δ), i.e., with $\delta_{ij}^\theta = 0$ for all $(i, j) \in \mathcal{E} \setminus \mathcal{E}_\Delta$. In addition, determine the largest ρ_Δ^θ ensuring robust stability of (10) (resp., (11) in discrete time).

Within this framework, it is possible to leverage Lem. 2.1 and yield the main theoretical contribution of this paper, represented by the robustness guarantees for system (10) when the target of a cyber-physical attack is a multitude of edges. In this direction, we study how the robust stability of (10) is affected by perturbations on all the weights $p_{uv}(\theta_{uv}) = w_{uv}$ attached to the connections $(u, v) \in \mathcal{E}_\Delta$ that are caused by the deviations of each subcodeword θ_{uv} .

As clarified later in more detail, the same three additional assumptions on the p_i 's proposed in [1] are *sufficient* to tackle Problem 3.1, namely, this robustness analysis is again restricted to a particular choice for the objective coding, that is for concave and Lipschitz continuous differentiable functions p_i . Thus, we let the i -th objective decoding function adopted in model (10) have the following properties.

Assumption 3.1: Each $p_i : \Theta \rightarrow \Pi_i$, with $i = 1, \dots, n$, has the subsequent characterization:

- (i) values $[p_i(\theta)]_j = p_{ij}(\theta_{ij})$, with $\theta_{ij} = [\theta_i]_j$, satisfy (5) for all $(i, j) \in \mathcal{E}$ and are not constant w.r.t. θ_{ij} ;
- (ii) p_{ij} is concave $\forall \theta \in \Theta$, i.e., $p_{ij}(\varsigma\eta_1 + (1-\varsigma)\eta_2) \geq \varsigma p_{ij}(\eta_1) + (1-\varsigma)p_{ij}(\eta_2)$, $\varsigma \in [0, 1]$, $\forall \eta_1, \eta_2 \in \Theta_{ij}$;
- (iii) p_{ij} is Lipschitz continuous and differentiable w.r.t. θ , implying $\exists K_{ij} \geq 0 : |p'_{ij}(\theta_{ij})| \leq K_{ij}$, $\forall (i, j) \in \mathcal{E}$.

Also, to provide analytical guarantees to multiple attacks striking the network, we will make use of the global quantity

$$K_{\Delta} = \max_{(u,v) \in \mathcal{E}_{\Delta}} \{K_{uv}\}. \quad (12)$$

B. Guarantees for multiple attacks

The guarantees in [1, Theorem IV.1] can be extended as follows for a continuous-time multiple-attack scenarios.

Theorem 3.1: Assume that the characterization for objective decoding functions p_i in Asm. 3.1 holds. For a structured injection attack $\delta^{\theta} \in \mathbf{\Delta}^{\theta}$ affecting all edges in \mathcal{E}_{Δ} define $\mathcal{R}_{\mathcal{E}_{\Delta}}(\mathcal{G})$ and K_{Δ} as in (4) and (12), respectively. Then the perturbed consensus protocol (10) is stable and achieves agreement for all δ^{θ} whenever

$$\|\delta^{\theta}\|_{\infty} < \rho_{\Delta}^{\theta} = (K_{\Delta} \mathcal{R}_{\mathcal{E}_{\Delta}}(\mathcal{G}))^{-1}, \quad (13)$$

independently from the values taken by any codeword $\theta \in \Theta$.

Proof: Similarly to the single edge attack case, Asm. 3.1 brings to each ordered logical step to conclude the thesis through Lem. 2.1. Indeed, (i)-(iii) lead to the determination of quantity $K_{ij}|\delta_{ij}^{\theta}|$, which can be seen as the maximum magnitude of an additive perturbation $\delta_{ij}^w := p_{ij}(\theta_{ij} + \delta_{ij}^{\theta}) - p_{ij}(\theta_{ij})$ affecting each p_{ij} , $\forall (i,j) \in \mathcal{E}$, independently from the transmitted codeword θ . Consequently, the fact that $|\delta_{uv}^w| \leq K_{uv}|\delta_{uv}^{\theta}|$ holds for each edge $(u,v) \in \mathcal{E}_{\Delta}$ implies that the following chain of inequalities is verified:

$$\|\delta^w\|_{\infty} \leq \max_{(u,v) \in \mathcal{E}_{\Delta}} \{K_{uv}|\delta_{uv}^{\theta}|\} \leq K_{\Delta} \|\delta^{\theta}\|_{\infty}. \quad (14)$$

Therefore, imposing inequality $K_{\Delta} \|\delta^{\theta}\|_{\infty} < \mathcal{R}_{\mathcal{E}_{\Delta}}^{-1}(\mathcal{G})$, in accordance with (4), leads to the thesis. ■

We notice that the leftmost inequality in (14) translates into an essential conservatism on the attack magnitude w.r.t. (4), similarly to the case $|\mathcal{E}_{\Delta}| = 1$. This can be modulated by choosing constants K_{ij} properly. Nonetheless, as soon as $|\mathcal{E}_{\Delta}| > 1$ holds, another essential conservatism that hinges on the attack's scale $|\mathcal{E}_{\Delta}|$ may arise w.r.t. its one-dimensional version. In the sequel, this concept is referred to as the \mathcal{E}_{Δ} -gap, or *resilience gap*, since it depends to the fact that, in general, one has $\mathcal{R}_{\mathcal{E}_{\Delta}}(\mathcal{G}) \geq \max_{(u,v) \in \mathcal{E}_{\Delta}} \{\mathcal{R}_{\{(u,v)\}}(\mathcal{G})\}$ (see discussion in Sec. III-C for more details).

Concerning, instead, the discrete-time guarantees for system (11) provided in [1, Theorem VI.1, Corollary VI.1], the following generalization can be made.

Corollary 3.1: Assume that the characterization for objective decoding functions p_i in Asm. 3.1 holds. Denote respectively with $\bar{w}_i = \sum_{j \in \mathcal{N}_i} |w_{ij}|$ and $\Psi_{\mathcal{G}} = \max_{i=1, \dots, n} \{\bar{w}_i\}$ the weighted degree of the i -th node and the maximum weighted degree of the underlying graph \mathcal{G} . Let a structured injection attack $\delta^{\theta} \in \mathbf{\Delta}^{\theta}$ affect all edges in \mathcal{E}_{Δ} and define

$$\psi_i(\delta_{uv}^{\theta}) = \bar{w}_i + K_{uv}|\delta_{uv}^{\theta}|, \quad \forall (u,v) \in \mathcal{E}_{\Delta}, \quad i = u, v. \quad (15)$$

Then the perturbed consensus protocol (11) is stable for all δ^{θ} such that both (13) and

$$\phi_{\mathcal{G}}(\delta^{\theta}) := \max \left\{ \Psi_{\mathcal{G}}, \max_{i \in \{u,v\}, (u,v) \in \mathcal{E}_{\Delta}} \psi_i(\delta_{uv}^{\theta}) \right\} < \epsilon^{-1} \quad (16)$$

hold for any fixed ϵ , independently from the values taken by any codeword $\theta \in \Theta$.

Proof: The proof is a generalization of that for Theorem VI.1 in [1], with the only difference that $\phi_{\mathcal{G}}(\delta^{\theta})$ in (16) is harder to be computed w.r.t. its single-edge-attack version $\phi_{\mathcal{G}}(\delta_{uv}^{\theta})$ because the combinatorial search space for the maximization of quantities $\psi_i(\delta_{uv}^{\theta})$ defined in (15) is, in general, larger in this kind of multiple-attack scenario. ■

Corollary 3.2: Under all the assumptions adopted in Cor. 3.1, defining K_{Δ} as in (12), $\mathcal{R}_{\mathcal{E}_{\Delta}}(\mathcal{G})$ as in (4) and setting $\epsilon < \Psi_{\mathcal{G}}^{-1}$, the perturbed consensus protocol (11) is stable for all δ^{θ} such that

$$\|\delta^{\theta}\|_{\infty} < \rho_{\Delta}^{\theta} = K_{\Delta}^{-1} \min\{\mathcal{R}_{\mathcal{E}_{\Delta}}^{-1}(\mathcal{G}), (\epsilon^{-1} - \Psi_{\mathcal{G}})\} \quad (17)$$

independently from the values taken by any codeword $\theta \in \Theta$. In particular, if ϵ is selected as follows

$$\epsilon \leq \epsilon_{\Delta}^* := (\Psi_{\mathcal{G}} + \mathcal{R}_{\mathcal{E}_{\Delta}}^{-1}(\mathcal{G}))^{-1} \quad (18)$$

then ρ_{Δ}^{θ} in (17) is maximized as ϵ varies and condition (13) needs to be fulfilled solely to guarantee robust stability.

Proof: Relation in (17) is the combined result of guarantee in (13) and that one obtainable imposing $\Psi_{\mathcal{G}} + K_{\Delta} \|\delta^{\theta}\|_{\infty} < \epsilon^{-1}$ to satisfy (16), since $\phi_{\mathcal{G}}(\delta^{\theta})$ can be upper bounded as $\phi_{\mathcal{G}}(\delta^{\theta}) \leq \Psi_{\mathcal{G}} + K_{\Delta} \|\delta^{\theta}\|_{\infty}$. On the other hand, relation (18) is derived enforcing $\mathcal{R}_{\mathcal{E}_{\Delta}}^{-1}(\mathcal{G}) \leq \epsilon^{-1} - \Psi_{\mathcal{G}}$ with the purpose to maximize ρ_{Δ}^{θ} as ϵ varies. ■

We conclude this subsection with the following remark.

Remark 3.1: Observe that no additional assumptions on the decoding functions p_i w.r.t. (i)-(iii) given in Subsec. III-A are required to solve Problem 3.1, i.e., to generalize the guarantees already yielded in [1] to a multiple-attack scenario.

C. Analysis of the resilience gap

In light of the previous theoretical results, we discuss here how decoding functions can be seen as a useful tool to compensate against the resilience gap (the \mathcal{E}_{Δ} -gap). This analysis starts by recalling the following preliminary proposition.

Proposition 3.1 ([16, Proposition V.3]):

For any weighted undirected graph \mathcal{G} it holds that

$$\mathcal{R}_{\mathcal{E}_{\Delta}}^*(\mathcal{G}) \leq \mathcal{R}_{\mathcal{E}_{\Delta}}(\mathcal{G}) \leq \mathcal{R}_{\mathcal{E}_{\Delta}}^{tot}(\mathcal{G}), \quad (19)$$

where $\mathcal{R}_{\mathcal{E}_{\Delta}}(\mathcal{G})$ is defined as in (4),

$$\mathcal{R}_{\mathcal{E}_{\Delta}}^*(\mathcal{G}) = \max_{(u,v) \in \mathcal{E}_{\Delta}} \{\mathcal{R}_{\{(u,v)\}}(\mathcal{G})\} \quad (20)$$

and $\mathcal{R}_{\mathcal{E}_{\Delta}}^{tot}(\mathcal{G}) = \text{tr}(P^{\top} R_{(\mathcal{T}, \mathcal{C})}^{\top} (R_{(\mathcal{T}, \mathcal{C})} W R_{(\mathcal{T}, \mathcal{C})}^{\top})^{-1} R_{(\mathcal{T}, \mathcal{C})} P)$.

Prop. 3.1 suggests us to define the \mathcal{E}_{Δ} -gap as

$$g(\mathcal{G}, \mathcal{E}_{\Delta}) = 1 - \mathcal{R}_{\mathcal{E}_{\Delta}}^*(\mathcal{G}) / \mathcal{R}_{\mathcal{E}_{\Delta}}(\mathcal{G}), \quad (21)$$

with $\mathcal{R}_{\mathcal{E}_{\Delta}}^*(\mathcal{G})$ and $\mathcal{R}_{\mathcal{E}_{\Delta}}(\mathcal{G})$ defined by (20) and (4). Indeed, the emerging conservatism related to the fact that multiple edges may be under attack ($|\mathcal{E}_{\Delta}| > 1$) grows as the value of $g(\mathcal{G}, \mathcal{E}_{\Delta}) \in [0, 1)$ increases. Also, exploiting (21), inequality (13) can be rewritten as

$$\|\delta^{\theta}\|_{\infty} < (1 - g(\mathcal{G}, \mathcal{E}_{\Delta})) / (K_{\Delta} \mathcal{R}_{\mathcal{E}_{\Delta}}^*(\mathcal{G})), \quad (22)$$

so that the quantity $\rho_{\Delta^*}^\theta := (K_\Delta \mathcal{R}_{\mathcal{E}_\Delta^*}(\mathcal{G}))^{-1}$ can be seen as the maximum value taken by ρ_Δ^θ in (13) as \mathcal{G} and \mathcal{E}_Δ vary. In particular, if $g(\mathcal{G}, \mathcal{E}_\Delta) = 0$ holds (indicating the absence of this kind of conservatism), there exists an edge $(u^*, v^*) \in \mathcal{E}_\Delta$ for which, by (22), the network \mathcal{G} is robustly stable for all perturbations δ^θ such that

$$\|\delta^\theta\|_\infty < (K_\Delta \mathcal{R}_{\{(u^*, v^*)\}}(\mathcal{G}))^{-1} = \rho_{\Delta^*}^\theta. \quad (23)$$

Remarkably, inequality (23) expresses the same guarantee provided by (13) as if $\mathcal{E}_\Delta = \{(u^*, v^*)\}$ was assigned, similarly to the one-dimensional case $|\mathcal{E}_\Delta| = 1$ debated in [1]. Since the case where only one edge is attacked clearly represents a scenario in which the conservatism due to $|\mathcal{E}_\Delta| > 1$ is lacking, if the \mathcal{E}_Δ -gap corresponding to a given setup $(\mathcal{G}, \mathcal{E}_\Delta)$ vanishes then the robustness of \mathcal{G} subject to perturbations δ^θ striking the subset \mathcal{E}_Δ is maximized. The latter observation leads us to wonder which are all the possible circumstances where $g(\mathcal{G}, \mathcal{E}_\Delta) = 0$ is satisfied? A partial yet extensive answer is given by the next proposition.

Proposition 3.2: Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{W})$, $\mathcal{E}_\Delta \subseteq \mathcal{E}$ the subset of perturbed edges in \mathcal{E} , and $g(\mathcal{G}, \mathcal{E}_\Delta)$ the \mathcal{E}_Δ -gap (21). Then $g(\mathcal{G}, \mathcal{E}_\Delta) = 0$ if either one of the following conditions hold: i) $|\mathcal{E}_\Delta| = 1$, or ii) $2 \leq |\mathcal{E}_\Delta| \leq n - 1 = |\mathcal{E}|$.

Proof: We prove each point of the statement separately.

i) If $|\mathcal{E}_\Delta| = 1$ then there exists only one edge $(u^*, v^*) \in \mathcal{E}_\Delta$ for which the mid inequality in (19) must hold tightly, since $\mathcal{R}_{\mathcal{E}_\Delta^*}(\mathcal{G}) = \mathcal{R}_{\mathcal{E}_\Delta}(\mathcal{G})$. It follows that $g(\mathcal{G}, \mathcal{E}_\Delta) = 0$.

ii) In this case, since \mathcal{G} is connected by assumption then it has to be a spanning tree; consequently, it can be also denoted as $\mathcal{G} = \mathcal{T}(\mathcal{E}, \mathcal{V}, \mathcal{W})$. Hence, one has $R_{(\mathcal{T}, \mathcal{C})} = I_{n-1}$ and, therefore, formula (4) yields

$$\mathcal{R}_{\mathcal{E}_\Delta}(\mathcal{T}) = \max_{k: e_k \in \mathcal{E}_\Delta} \{|w_k^{-1}|\}, \quad \forall \mathcal{E}_\Delta \subseteq \mathcal{E}. \quad (24)$$

Observe that (24) also holds for subsets \mathcal{E}_Δ such that $|\mathcal{E}_\Delta| = 1$ and the max over $\{k: e_k \in \mathcal{E}_\Delta\}$ in (24) is taken exactly as in (20). We conclude that $g(\mathcal{G}, \mathcal{E}_\Delta) = 0$ occurs even for $|\mathcal{E}_\Delta| \geq 2$ whenever $\mathcal{G} = \mathcal{T}(\mathcal{E}, \mathcal{V}, \mathcal{W})$, as the mid inequality in (19) is tightly satisfied under these conditions. ■

To conclude, we note that the \mathcal{E}_Δ -gap in (22) can be mitigated by an appropriate re-design procedure of K_Δ .

Proposition 3.3: Denote with $\rho_{\Delta^*}^\theta(K_\Delta)$ the value of $\rho_{\Delta^*}^\theta$ as K_Δ varies. Set $K'_\Delta := (1 - g(\mathcal{G}, \mathcal{E}_\Delta))K_\Delta$. Then

$$\rho_{\Delta^*}^\theta(K'_\Delta) > \rho_{\Delta^*}^\theta(K_\Delta), \quad \forall K_\Delta > 0.$$

Prop. 3.3 introduces an interesting tradeoff, where conservatism is reduced at the expense of shrinking the image of the decoding function (see also Fact IV.1 in [1]). This leads to an increasing demand of the encryption capabilities.

IV. NUMERICAL EXAMPLES

We here focus on a potential application of the proposed technique to semi-autonomous networks (SANs) consisting of leader-follower autonomous agents [18]. Let us assume that subset $\mathcal{V}_l \subseteq \mathcal{V}$, $\mathcal{V}_l \neq \emptyset$, collects all the leader agents of a SAN $\mathcal{S} = (\mathcal{G}, \mathcal{U})$, where $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{W})$ is an undirected and connected graph representing the network interactions of \mathcal{S} and $\mathcal{U} = \{\mathbf{u}_1(t), \dots, \mathbf{u}_{|\mathcal{V}_l|}(t)\}$, $\mathbf{u}_\ell(t) \in \mathbb{R}^D$, denotes

the set of external inputs that directly influence each leader agents in \mathcal{V}_l . Typically, setting $\mathbf{u}(t) = \text{vec}_{i \in \mathcal{V}_l}(\mathbf{u}_i(t))$, the continuous-time dynamics for SANs are yielded by

$$\dot{\mathbf{x}} = (L_B(\mathcal{G}) \otimes I_D)\mathbf{x} + (B \otimes I_D)\mathbf{u}, \quad (25)$$

where $L_B(\mathcal{G}) = L(\mathcal{G}) + \text{diag}(B\mathbf{1}_{|\mathcal{V}_l|})$, with $B \in \mathbb{R}^{n \times |\mathcal{V}_l|}$ such that $[B]_{i\ell} > 0$, if agent i belongs to the leader set \mathcal{V}_l ; $[B]_{i\ell} = 0$, otherwise. Also, a discrete-time version of (25) can be evidently provided by

$$\mathbf{x}(t+1) = (I_N - \epsilon(L_B(\mathcal{G}) \otimes I_D))\mathbf{x}(t) + (B \otimes I_D)\mathbf{u}(t), \quad (26)$$

where the quantity ϵ preserves the same meaning of (11). The stability of SANs endowed with the discussed security mechanisms can be analyzed by observing that the positive semi-definiteness of $L(\mathcal{G})$ clearly implies the positive semi-definiteness of $L_B(\mathcal{G})$. Consequently, it is sufficient to ensure Thm. 3.1 and Cor. 3.1 to hold in order to respectively guarantee the robust stability of protocols (25) and (26).

In the following lines, we examine the SAN $\mathcal{S} = (\mathcal{G}, \mathcal{U})$ whose interconnections are modeled through the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{W})$ illustrated in Fig. 1(a): the sole leader (node 1) is marked in orange, the followers (remaining nodes) are marked in light blue, and the edge weights $\mathcal{W} = \{w_k\}_{k=1}^5$ given by $w_1 = w_{12}$, $w_2 = w_{35}$, $w_3 = w_{46}$, $w_4 = w_{24}$, $w_5 = w_{23}$ are chosen according to the same picture. Moreover, we set $D = 1$, $[B]_{i\ell} = 1$ for $\ell = 1$ if $i = 1$, and $\mathcal{U} = \{u_1\}$, with $u_1 = -0.5$. We test the proposed security methods both in continuous and discrete time by uniformly adopting for all the edges the following objective decoding function

$$p_\Delta(\eta) = \begin{cases} K_\Delta \left(\frac{4}{13} \sqrt{\eta + 1} + 1 \right), & \text{if } \eta \geq 3; \\ K_\Delta \left(-\frac{2}{13} \eta^2 + \eta \right), & \text{if } 0 \leq \eta < 3; \\ K_\Delta \eta, & \text{if } \eta < 0; \end{cases} \quad (27)$$

where $K_\Delta > 0$ is a suitable Lipschitz constant for $p_\Delta(\eta)$, to be selected in order to ensure that $\|\delta^w\|_\infty \leq \rho_\Delta^\theta = 0.5$. More in detail, we consider two kinds of malicious attacks hitting part of \mathcal{W} . The first external perturbation is delivered against edge (1,2) (red strike in Fig. 1(a)). Whereas, the second one involves not only edge (1,2) but also edges (3,5) and (4,6) (yellow strikes in Fig. 1(a)). Hence, under this specific adversarial setup, we denote the set of edges under attack as $\mathcal{E}_1 = \{(1,2)\}$ and $\mathcal{E}_2 = \{(1,2), (3,5), (4,6)\}$ for the first and second kind of perturbation, respectively.

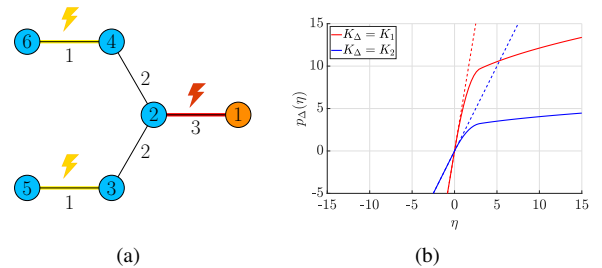
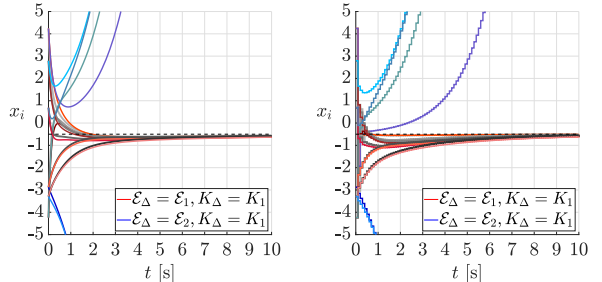


Fig. 1: (a): underlying topology of the given SAN (leaders in orange and followers in light blue); (b): Objective decoding functions considered for this case study.

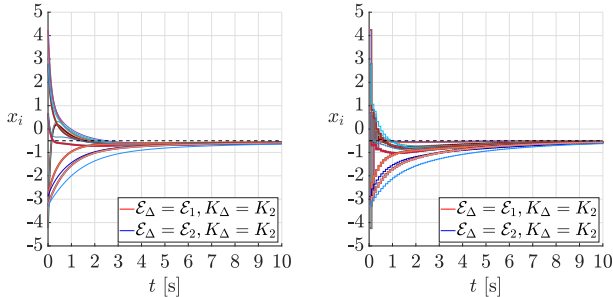
Observing that \mathcal{G} in Fig. 1(a) is an undirected tree (i.e. $\mathcal{G} = \mathcal{T}$) then $\mathcal{R}_{\mathcal{E}_\Delta}(\mathcal{G})$ can be computed as in (24). One can set $\epsilon = (\mathcal{R}_{\mathcal{E}_1}^{-1}(\mathcal{G}) + \Psi_{\mathcal{G}})^{-1} = (2w_1 + w_4 + w_5)^{-1} = 0.1$ in order to satisfy (18) for both $\mathcal{E}_1, \mathcal{E}_2$; thus, the Lipschitz constant K_Δ of p_Δ can be finally designed according to (24) and (13) so that it respectively takes values $K_1 = 6$ and $K_2 = 2$ for perturbations on edges \mathcal{E}_1 and \mathcal{E}_2 . In Fig. 1(b), the decoding function (27) is then depicted (solid curves) for such values of K_Δ along with the corresponding (dashed) lines having slope equal to its Lipschitz constant.

We now discuss the robust stability of the SAN under investigation. In particular, we focus on the trajectories of (25) and (26) for \mathcal{S} subject to the following perturbations: $\delta_1^\theta = -0.5\rho_\Delta^\theta [10000]^\top$, $\delta_2^\theta = -0.5\rho_\Delta^\theta [11100]^\top$, where δ_1^θ strikes the edge in \mathcal{E}_1 and δ_2^θ strikes the three edges in \mathcal{E}_2 . It is crucial to note that if $K_\Delta = K_1$ is selected as in the single-case attack scenario proposed in [1] adopting p_Δ to counter the perturbation δ_2^θ , then no stability guarantees can be given (see Figs. 2(a)-2(b)). Indeed, despite this choice allows to mitigate the malicious effects of δ_1^θ , the additional perturbations on edges (3, 5) and (4, 6) are capable of disrupting the network state convergence towards u_1 as $t \rightarrow +\infty$ because $K_\Delta = K_1$ does not match condition (13). Instead, Figs. 2(c)-2(d) show that such an issue is overtaken if $K_\Delta = K_2$ is set ensuring proper state convergence in accordance with (13).

In light of this, we can appreciate that the extension of the SBDC protocol towards multiple-attack scenarios requires, in general, more encryption capabilities and attention to the design phase in order to ensure robust stability compared to the single-attack case. Indeed, in this simulation, we have seen that only by considering $\mathcal{R}_{\mathcal{E}_\Delta}(\mathcal{G})$ computed as in (4) and lowering the value for K_Δ the agreement can be reached.



(a) Divergence (continuous time). (b) Divergence (discrete time).



(c) Convergence (continuous time). (d) Convergence (discrete time).

Fig. 2: Behaviors of the SAN trajectories as \mathcal{E}_Δ and p_Δ vary. According to ϵ , each discrete time stamp is set to 0.1 s. (a)-(b): stability guarantees for attack δ_1^θ on \mathcal{E}_1 but not for δ_2^θ on \mathcal{E}_2 ; (c)-(d): stability guarantees for the attacks δ_i^θ on \mathcal{E}_i , $i = 1, 2$.

V. FINAL REMARKS AND FUTURE DIRECTIONS

The paper broadens the approach devised in [1] for secure consensus networks to a multiple attack scenario, both in the continuous and discrete time domains. Even in this framework, small-gain-theorem-based stability guarantees are yielded to this aim and, remarkably, no additional assumption is needed to provide such a generalization. In addition, the conservatism arising from a multiplicity of threats has been addressed and analyzed. Future works will involve new developments towards other multi-agent protocols, such as nonlinear consensus and distance-based formation control.

REFERENCES

- [1] M. Fabris and D. Zelazo, "Secure consensus via objective coding: Robustness analysis to channel tampering," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1–13, 2022.
- [2] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, Jan 2007.
- [3] M. Meng, G. Xiao, and B. Li, "Adaptive consensus for heterogeneous multi-agent systems under sensor and actuator attacks," *Automatica*, vol. 122, p. 109242, 2020.
- [4] J. Wang, X. Deng, J. Guo, and Z. Zeng, "Resilient consensus control for multi-agent systems: A comparative survey," *Sensors*, vol. 23, no. 6, 2023.
- [5] D. Zhang, G. Feng, Y. Shi, and D. Srinivasan, "Physical safety and cyber security analysis of multi-agent systems: A survey of recent advances," *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 2, pp. 319–333, 2021.
- [6] Y. Yang, Y. Xiao, and T. Li, "Attacks on formation control for multiagent systems," *IEEE Transactions on Cybernetics*, pp. 1–13, 2021.
- [7] S. Huo, H. Wu, and Y. Zhang, "Secure consensus control for multi-agent systems against attacks on actuators and sensors," *International Journal of Robust and Nonlinear Control*, vol. 32, no. 8, pp. 4861–4877, 2022.
- [8] Z.-W. Liu, Y.-L. Shi, H. Yan, B.-X. Han, and Z.-H. Guan, "Secure consensus of multi-agent systems via impulsive control subject to deception attacks," *IEEE Transactions on Circuits and Systems II: Express Briefs*, pp. 1–1, 2022.
- [9] A.-A. Gao, A.-H. Hu, and Z.-X. Jiang, "Secure consensus of multi-agent systems based on event-triggered impulsive control," *Journal of Computer Applications*, p. 0, 2022.
- [10] S. M. Elkhider, S. El-Ferik, and A.-W. A. Saif, "Denial of service attack of qos-based control of multi-agent systems," *Applied Sciences*, vol. 12, no. 9, 2022.
- [11] J. Wang, Y. Li, Z. Duan, and J. Zeng, "A fully distributed robust secure consensus protocol for linear multi-agent systems," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 69, no. 7, pp. 3264–3268, 2022.
- [12] M. Sathishkumar and Y.-C. Liu, "Resilient memory event-triggered consensus control for multi-agent systems with aperiodic dos attacks," *International Journal of Control, Automation and Systems*, pp. 1–14, 2022.
- [13] C. Gao, X. He, H. Dong, H. Liu, and G. Lyu, "A survey on fault-tolerant consensus control of multi-agent systems: trends, methodologies and prospects," *International Journal of Systems Science*, vol. 53, no. 13, pp. 2800–2813, 2022.
- [14] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Systems & Control Letters*, vol. 53, no. 1, pp. 65–78, 2004.
- [15] J. Lunze, *Networked control of multi-agent systems*. Edition MoRa, 2019.
- [16] D. Zelazo and M. Bürger, "On the robustness of uncertain consensus networks," *IEEE Transactions on Control of Network Systems*, vol. 4, no. 2, pp. 170–178, 2017.
- [17] M. Fabris, G. Michieletto, and A. Cenedese, "A general regularized distributed solution for system state estimation from relative measurements," *IEEE Control Systems Letters*, vol. 6, pp. 1580–1585, 2022.
- [18] A. Chapman, *Semi-Autonomous Networks: Effective Control of Networked Systems through Protocols, Design, and Modeling*. Springer, 2015.