Northumbria Research Link

Northumbria
University
NEWCASTLE

Resilient Consensus for Robust Multiplex Networks with Asymmetric Confidence Intervals

Yilun Shang

Abstract—**The consensus problem with asymmetric confidence intervals considered in this paper is characterized by the fact that each agent can have optimistic and/or pessimistic interactions with its neighbors. To deal with the asymmetric confidence scenarios, we introduce a novel multiplex network presentation for directed graphs and its associated connectivity concepts including the pseudo-strongly connectivity and graph robustness, which provide a resilience characterization in the presence of malicious nodes. We develop distributed resilient consensus strategies for a group of dynamical agents with locally bounded Byzantine agents in both continuous-time and discrete-time multi-agent systems. Drawing on our multiplex network framework, much milder connectivity conditions compared to existing works are proposed to ensure resilient consensus. The results are further extended to cope with resilient scaled consensus problems which allow both cooperative and antagonistic agreements among agents. Numerical examples are also exhibited to confirm the theoretical results and reveal the factors that affect the speed of convergence in our multiplex network framework.**

Index Terms—**Consensus, robust multiplex network, multi-agent system, asymmetric interaction.**

I. Introduction

OVER the past decade, consensus algorithms have attracted an ongoing attention and become a significant building block of various distributed systems and control protocols. In a standard consensus problem, all agents seek to reach an agreement by updating their states based on local interactions with their neighbors in continuous and discrete time [1]. Connectivity of the underlying communication topology arguably is central to the coordination of multi-agent systems, where agent interactions form an interconnected network. A variety of consensus conditions, such as strongly connectedness and jointly connectivity, have been reported for multi-agent systems under undirected or directed as well as fixed or time-dependent topologies etc. [2]–[4].

In many consensus-based applications, agents do not always operate in a benign environment. It is likely that one or more agents are compromised due to faults like hardware failures or malicious attacks in cyber-physical systems and social networks [5]–[7]. Such misbehaving agent requires additional connectivity of the network to cope with consensus reaching objective. A notion of r-robustness is introduced in [8], [9] to facilitate asymptotic distributed consensus in discrete-time multi-agent systems. In particular, based upon the Weighted-Mean Subsequence Reduced (W-MSR) protocol, a $(2r+1)$-robust network is able to withstand r malicious agents presenting in the neighborhood of any agent while still achieve consensus among the normal cooperative agents. The scheme has been later extended to tackle more complicated agent dynamics such as higher-order [10], hybrid dynamics [11], [12], and switching dynamics with heterogeneous agents [13]. Resilient quantized consensus problem has been solved in [14] for discrete-time multi-agent systems, where the states of agents are confined to integers. Resilient interval consensus and resilient group consensus problems have been initiated in [15] and [16], respectively, taking into consideration of state constraints as well as topological clustering. W-MSR algorithms have also been applied in [17] to tackle distributed state estimation for linear time-invariant systems over directed acyclic graphs.

Despite remarkable merits (such as low complexity and distributed update rule) of the W-MSR strategies, the high connectivity requirement on the communication topology limits their applications as a dense network may be costly. To achieve structural robustness without adding extra links, trusted nodes are introduced in [18]. These nodes will be protected from attack at all times so that resilient consensus can be achieved under milder connectivity. A similar edge-oriented version is examined in [19], where trusted links are proposed to improve network robustness and achieve convergence in finite time. Another effort along this line is to endow agents with diverse vulnerabilities or weaknesses [20], where a malicious agent can only attack a particular type of weakness, and hence consensus may be achieved in a relatively sparse network.

In the above mentioned literature on resilient consensus, all agents are assumed to take in neighbors' state values which are both larger and smaller than its own value, or with symmetric confidence intervals in the language of opinion dynamics [21]. However, in some real-life scenarios, agents interaction is not entirely two-sided. Examples include optimistic/pessimistic agents in opinion models [22], where optimistic one only accepts higher opinions while pessimistic one lower opinions, and unilateral gossip-based algorithms for distributed systems [23]. The idea of asymmetric confidence intervals has been explored in studying consensus for cooperative monotone systems of continuous-time agents [24].

Here, we develop a novel distributed coordination framework to reduce structural robustness requirement in resilient consensus problems [8]. Our approach builds on the asymmetric confidence intervals and provides resilience against locally bounded malicious agents in both continuous-time and discrete-time multi-agent systems. The contributions of the work are as follows. Firstly, we introduce a new topology

Y. Shang is with the Department of Computer and Information Sciences, Faculty of Engineering and Environment, Northumbria University, Newcastle upon Tyne, NE1 8ST, UK (e-mail: shylmath@hotmail.com).

Manuscript received Jan xx, 2019; revised April xx, 2020.

property of network robustness over a multiplex network, which is induced by the agent dynamics with asymmetric confidence intervals. This multiplex network robustness bridges the key gap between resilient consensus over simplex network and multiplex network, and extends the notion of robustness for simplex graphs underpinning existing W-MSR algorithms [9], [11]–[13]. Secondly, based on the above graph-theoretic property we introduce a connectivity concept, referred to as pseudo-strong connectivity, which unifies strong connectivity and spanning tree condition in directed graphs. This connectivity notion is tailored for consensus-reaching over multiplex network structure induced by asymmetric confidence intervals and provides a means for graph-theoretic study of multiplex networks. Thirdly, we develop resilient consensus strategies for both continuous-time and discrete-time cooperative agents in the presence of Byzantine agents, who are anonymous and have complete knowledge of the network. Leveraging on the pseudo-strong connectivity, weak conditions are established to guarantee consensus when the number of Byzantine agents remain locally bounded. Unlike the previous works [18]–[20], no extra assumptions like trusted nodes or diverse vulnerabilities are imposed. Finally, we extend our resilient consensus framework to encompass resilient scaled consensus problems, where an agent may reach individual value positive/negative-proportionally associated with other agents offering further flexibility. Our numerical simulations unravel the factors that play a role on convergence speed and underscore the effectiveness of the multiplex framework, which facilitates resilient consensus under milder connectivity conditions.

The rest of the paper is organized as follows. Section II presents some graph theoretical preliminaries and sets up the resilient consensus models. Convergence analysis is carried out for continuous-time and discrete-time dynamics in Sections III and IV, respectively. Numerical examples are provided in Section V. Finally, concluding remarks are given in Section VI with discussions regarding alternative solutions to cope with asymmetric confidence scenarios.

II. Problem formulation

A. Graph theory

A directed multiplex network $G(V, E^o, E^p)$ consists of a node set $V = \{1, 2, \cdots, n\}$ and two layers: the optimistic layer $G^o(V, E^o)$ and the pessimistic layer $G^p(V, E^p)$ [25], [36]. An edge (i, j) in E^o (or E^p) is an edge from i to j in the optimistic (or pessimistic) layer. Denote the overall edge set of G by $E = E^o \cup E^p$ and the corresponding adjacency matrix by $A = (a_{ij}) \in \mathbb{R}^{n\times n}$, where $a_{ij} > 0$ when $(j, i) \in E$ and $a_{ij} = 0$ otherwise. We will also write $G(V, E) = G^o \cup G^p$ as a single network combining the two types of edges. Nodes in $V = C \cup B$ can be divided into cooperative nodes in C, which adopt our resilient consensus strategies detailed in Section II.A and Byzantine nodes in B defined in Definition 6 below. The (in-degree) neighborhood of a node i can be defined as $N_i = N_i^o \cup N_i^p$, where $N_i^o = N_i^o(G) := \{j \in V : (j, i) \in E^o\}$ contains optimistic neighbors and $N_i^p = N_i^p(G) := \{j \in V : (j, i) \in E^p\}$ contains pessimistic neighbors. For a node $j \in V$, it is possible that $j \in N_i^o \cap N_i^p$; see Definition 5 for a precise definition of these two layers. A finite sequence $i_1, i_2, \cdots, i_k$ of different nodes with $(i_l, i_{l+1}) \in E$ for $1 \le l \le k-1$ is called a directed path with length n in $G(V, E)$. $G(V, E)$ is said to contain a spanning tree if there is a node v (called root node), from which any other node in V can be reached via a directed path. If every node can be a root, then the graph is called strongly connected.

Multiplex networks as an important example of multi-layer complex networks [26] have become a prominent research field in network science in recent years. They are a major tool for the analysis of interacting components in real networked systems with different types of interactions. A diverse range of applications have been found in, for example, synchronization problems [27], interbank market [28], social contagions [29], cognitive psychology [30], molecular biology [31], and community detection [32], to name just a few. Structural properties of multiplex networks, including core [33], communicability [34], and isomorphism [35], have attracted much attention from the perspective of network theory and often represent an essential extension to their counterparts in classical graph theory.

Let $|\cdot|$ be the size of a given set. We next introduce the reachability and robustness notions for our multiplex network $G(V, E^o, E^p)$.

Definition 1 (r-reachable set). A set $S \subseteq V$ is called r-reachable in $G(V, E^o, E^p)$ if there exist $i, j \in S$ (possibly $i = j$) such that $|N_i^o \backslash S| \ge r-1$, $|N_j^p \backslash S| \ge r-1$, and $\max\{|N_i^o \backslash S|, |N_j^p \backslash S|\} \ge r$.

Definition 2 (r-robust multiplex graph). $G(V, E^o, E^p)$ is called r-robust if for any nonempty and disjoint $S_1, S_2 \subseteq V$ at least one of them is r-reachable.

It is easy to see if $E^o = E^p$ then r-robustness of $G(V, E^o, E^p)$ is equivalent to r-robustness of $G(V, E)$ [8]. The following result shows that robustness is a monotone property with respect to edge deletion.

Lemma 1. *If $G(V, E^o, E^p)$ is r-robust and G' is obtained from G by each node removing up to $s < r$ incoming edges in E^o and up to $s < r$ incoming edges in E^p, then G' is $(r-s)$-robust.*

Proof. For any nonempty subsets $S_1 \cap S_2 = \emptyset$, there must be one of them, say S_1, being r-reachable. Therefore, we have some nodes $i, j \in S_1$ satisfying $|N_i^o(G)\backslash S_1| \ge r-1$, $|N_j^p(G)\backslash S_1| \ge r-1$, and $\max\{|N_i^o(G)\backslash S_1|, |N_j^p(G)\backslash S_1|\} \ge r$. In light of the edge removal procedure, we conclude $|N_i^o(G')\backslash S_1| \ge r-s-1$, $|N_j^p(G')\backslash S_1| \ge r-s-1$, and $\max\{|N_i^o(G')\backslash S_1|, |N_j^p(G')\backslash S_1|\} \ge r-s$. This means G' is $(r-s)$-robust. □

Albeit varied connectivity concepts in graph and network science, we propose the following novel connectivity property for multiplex networks.

Definition 3 (pseudo-strong connectivity). $G(V, E^o, E^p)$ is called pseudo-strongly connected if for any pair $(i, j) \in V \times V$ (with $i \ne j$), there exists a node $k_{ij} \in V$ (possibly $k_{ij} = i$ or $k_{ij} = j$) with a directed path from k_{ij} to i and another directed path from k_{ij} to j in $G(V, E)$.

Given a pair of nodes i and j, if there is a directed path from i to j, we can choose $k_{ij} = i$ to satisfy Definition 3.

Here, $k_{ij} = i$ is thought of as having a path of length "zero" from k_{ij} to i.

Lemma 2. *$G(V, E^o, E^p)$ is 1-robust if and only if it is pseudo-strongly connected.*

Proof. (Sufficiency) We will prove by contradiction. Suppose $G(V, E^o, E^p)$ is not 1-robust but it is pseudo-strongly connected. There must exist two nonempty subsets with $S_1 \cap S_2 = \emptyset$ such that none of them have any in-degree neighbors outside of themselves in $G^o(V, E^o)$ or $G^p(V, E^p)$. We take $i \in S_1$ and $j \in S_2$ so that the node k_{ij} in Definition 3 cannot be found. This contradicts pseudo-strong connectivity.

(Necessity) We first show that $G(V, E)$ has a spanning tree. In fact, if this does not hold, then we consider the decomposition of its strongly connected components. Since $G(V, E)$ has no spanning tree, there are two strongly connected components S_1 and S_2 having no incoming edges from outside of themselves in $G(V, E)$. This contradicts 1-robustness. Hence, we proved that $G(V, E)$ admits a spanning tree.

Next, we show $G(V, E^o, E^p)$ is pseudo-strongly connected. Since $G(V, E)$ has a spanning tree, for any pair $(i, j) \in V \times V$ with $i \neq j$ we take k_{ij} as the root node of the spanning tree. Thus, there is a directed path from k_{ij} to i and another directed path from k_{ij} to j following the spanning tree. □

Remark 1. The pseudo-strong connectivity concept can be viewed as a generalization of strongly connectivity and spanning tree condition in classical graph theory (i.e., simplex graphs). (1) It is not difficult to see when $E^o = \emptyset$ (or $E^p = \emptyset$), $G(V, E^o, E^p)$ is pseudo-strongly connected if and only if $G^p(V, E^p)$ (or $G^o(V, E^o)$) is strongly connected. (2) When $E^o = E^p = E$, $G(V, E^o, E^p)$ is pseudo-strongly connected if and only if $G(V, E)$ has a spanning tree. In fact, if $G(V, E)$ contains a spanning tree, by the proof of Lemma 2 we see that $G(V, E^o, E^p)$ is pseudo-strongly connected. The other direction can be shown by contradiction. Suppose that $G(V, E)$ has no spanning tree but $G(V, E^o, E^p)$ is pseudo-strongly connected. Hence, in the decomposition of strongly connected components of $G(V, E)$, there must be two components S_1 and S_2 having no incoming edges from outside of themselves. We choose $i \in S_1$ and $j \in S_2$. These two nodes have no common ancestor in $G(V, E)$. This contradicts Definition 3.

B. Resilient consensus strategies

In our resilient consensus model, we start by describing the definition of consensus over the node set $V = \{1, 2, \cdots, n\} = C \cup B$. The state of each agent $i \in V$ at time $t \geq 0$ is given by $x_i(t) \in \mathbb{R}$. We are interested in reaching consensus among the cooperative agents in C.

Definition 4 (resilient consensus). The cooperative agents in $G(V, E)$ are said to achieve resilient consensus if, for any initial configuration $\{x_i(0)\}_{i \in V}$, there exists $y \in \mathbb{R}$ such that $\lim_{t\to\infty} x_i(t) = y$ for all $i \in C$.

When $B = \emptyset$, resilient consensus problems become the standard consensus problems with only cooperative agents [1]. Any cooperative agent i in C follows a general dynamical system given by

$$\dot{x}_i(t) = f_i\left(\{t, x_j^i(t) : j \in N_i \cup \{i\}\}\right) \tag{1}$$

in continuous time, or

$$x_i(t+1) = f_i\left(\{t, x_j^i(t) : j \in N_i \cup \{i\}\}\right) \tag{2}$$

in discrete time. Here, the notation $x_j^i(t) \in \mathbb{R}$ represents the value sent from node j to node i. If $j \in C$, we assume $x_j^i(t) = x_j(t)$, i.e., cooperative agents always transmit their true states.

To specify the multiplex structure of $G(V, E^o, E^p)$ induced by the asymmetric confidence intervals, we define the optimistic layer $G^o(V, E^o)$ and pessimistic layer $G^p(V, E^p)$ as follows.

Definition 5 (optimistic and pessimistic edges). Given $i \in C$ and $j \in N_i$, the edge $(j, i) \in E^o$ if for any $b_j > b_i$ there is $\varepsilon > 0$ such that

$$f_i\big(\{t, x_j = b_j, x_l = b_i : l \in (N_i \cup \{i\}) \backslash \{j\}\}\big) \geq \varepsilon(b_j - b_i), \quad \forall t \geq 0.$$

Similarly, the edge $(j, i) \in E^p$ if for any $b_j < b_i$ there is $\varepsilon > 0$ such that

$$f_i\big(\{t, x_j = b_j, x_l = b_i : l \in (N_i \cup \{i\}) \backslash \{j\}\}\big) \leq \varepsilon(b_j - b_i), \quad \forall t \geq 0.$$

Clearly, for a cooperative node i if (j, i) is an optimistic (or pessimistic) edge, i is willing to update its states on the basis of the input from j, which has a higher (or lower) value. The idea is similar to the bicolored graph examined in [22] but the multiplex topology was not explicitly defined like this. Next, we consider the Byzantine agents in B, which can apply totally different update rules with the aim to corrupt the multi-agent system.

Definition 6 (Byzantine agents). An agent $i \in B$ is called Byzantine if it exerts a different update strategy $\hat{f}_i$ from those of cooperative agents in (1)-(2), or it sends different values to different neighbors at some time $t > 0$.

Byzantine nodes are assumed to be able to collude with other Byzantine nodes and potentially have complete knowledge of the whole network [11], [12], [18], [20]. Therefore, a Byzantine node i can apply arbitrary control strategies and freely assign an edge (j, i) with $j \in N_i$ to E^o or E^p. While cooperative agents have no knowledge about the identity and number of Byzantine ones, it is not unreasonable to assume that the number of Byzantine agents has an upper bound in an in-degree neighborhood. Specifically, we assume that $|N_i \cap B| \leq r$ for any $i \in C$. This condition is often referred to as the locally bounded assumption; see e.g. [8]. In other words, a cooperative agent is not required to have the capacity to determine which neighbor is Byzantine or not; all that it knows is an upper bound r.

Given the parameter r, we propose the following resilient consensus strategy for continuous-time dynamical agents (1) as follows. At time t, each cooperative agent $i \in C$ receives the state values $\{x_j^i(t)\}_{j \in N_i}$ from its neighbors and determines the E^o and E^p layers for its incoming edges. Sort respectively the two lists $\{x_j^i(t)\}_{j \in N_i^o}$ and $\{x_j^i(t)\}_{j \in N_i^p}$ in a descending manner. Remove up to r highest values in $\{x_j^i(t)\}_{j \in N_i^o}$ which are strictly larger than $x_i(t)$, and up to r lowest values in $\{x_j^i(t)\}_{j \in N_i^p}$ which are strictly smaller than $x_i(t)$. All removed indices are recorded in a set $R_i(t)$ (see the Removal Algorithm 1 given below). We instantiate f_i in (1) for $t \geq 0$ as

$$\dot{x}_i(t) = \sum_{j \in (N_i \cup \{i\}) \setminus R_i(t)} a_{ij} \rho_{ij}(t) f_{ij}\left(x_j^i(t), x_i(t)\right), \quad (3)$$

where $\rho_{ij}(t) > 0$, $f_{ij}(b_i, b_j)$ is assumed to be a locally Lipschitz continuous function, $f_{ij}(b_i, b_j) = 0 \Leftrightarrow b_i = b_j$, and $(b_i - b_j) f_{ij}(b_i, b_j) > 0$ when $b_i \neq b_j$. Moreover, $\lim_{t \to \infty} \rho_{ij}(t) = 0$ if (i) $(j,i) \in E^o \setminus E^p$ and $b_i < b_j$ or (ii) $(j,i) \in E^p \setminus E^o$ and $b_i > b_j$; $\rho_{ij}(t) \equiv 1$ otherwise.

Similarly, for discrete-time dynamical agents (2) we adopt an analogous censoring procedure and instantiate f_i in (2) for $t \geq 0$ as

$$x_i(t+1) = \sum_{j \in (N_i \cup \{i\}) \setminus R_i(t)} w_{ij}(t) x_j^i(t), \quad (4)$$

where $w_{ij}(t) = 0$ when $j \notin (N_i \cup \{i\}) \setminus R_i(t)$; $w_{ij}(t) > 0$ but $\lim_{t \to \infty} w_{ij}(t) = 0$ if $j \in \Theta_{i1}(t) \cup \Theta_{i2}(t)$ with $\Theta_{i1}(t) := \{j : (j,i) \in E^o \setminus E^p, x_j^i(t) < x_i(t)\}$ and $\Theta_{i2}(t) := \{j : (j,i) \in E^p \setminus E^o, x_j^i(t) > x_i(t)\}$; and $w_{ij}(t) \geq w > 0$ for some constant $w > 0$ when $j \in ((N_i \cup \{i\}) \setminus R_i(t)) \setminus (\Theta_{i1}(t) \cup \Theta_{i2}(t))$. Moreover, we assume $\sum_{j \in (N_i \cup \{i\}) \setminus R_i(t)} w_{ij}(t) = 1$.

Removal Algorithm 1 for a cooperative agent $i \in C$

Input: $x_i(t)$, $\{x_j^i(t)\}_{j \in N_i^o}$, $\{x_j^i(t)\}_{j \in N_i^p}$
Output: $R_i(t)$
sort $\{x_j^i(t)\}_{j \in N_i^o}$ decreasingly as
  $L^o := \{x_{j_1}^i(t), x_{j_2}^i(t), \cdots, x_{j_{|N_i^o|}}^i(t)\}$
let $R_i(t) = \emptyset$, $\bar{\ell} = \max\{1 \leq l \leq |N_i^o| : x_{j_l}^i(t) > x_i(t)\}$
if $\bar{\ell} > r$
  $\bar{\ell} = r$
end if
for $l = 1$ till $l = \bar{\ell}$
  add v_{j_l} into $R_i(t)$
end for
sort $\{x_j^i(t)\}_{j \in N_i^p}$ decreasingly as
  $L^p := \{x_{j_1}^i(t), x_{j_2}^i(t), \cdots, x_{j_{|N_i^p|}}^i(t)\}$
let $\underline{\ell} = \min\{1 \leq l \leq |N_i^p| : x_{j_l}^i(t) < x_i(t)\}$
if $|N_i^p| - \underline{\ell} + 1 > r$
  $\underline{\ell} = |N_i^p| - r + 1$
end if
for $l = \underline{\ell}$ till $l = |N_i^p|$
  add v_{j_l} into $R_i(t)$
end for

Remark 2. The proposed strategies can be viewed as generalizations of classical W-MSR algorithms [8]–[11] by accommodating asymmetric confidence intervals as well as non-autonomous dynamics. A common choice of f_{ij} in continuous time in (3) can be $f_{ij}(b_i, b_j) = b_i - b_j$ [1]. The assumption on $\rho_{ij}(t)$ implies that if the edge (j,i) is optimistic (but not pessimistic) then lower states of node j will not be able to influence node i in the long run. Likewise, if (j,i) is pessimistic (but not optimistic) then higher states of node j will not influence node i in the long run. For a discrete-time system, the adaptive weight $w_{ij}(t)$ in (4) can be typically taken as $w_{ij}(t) = (|(N_i \cup \{i\}) \setminus R_i(t)| - |\Theta_{i1}(t) \cup \Theta_{i2}(t)|)^{-1} - \delta(t)$ for $j \in ((N_i \cup \{i\}) \setminus R_i(t)) \setminus (\Theta_{i1}(t) \cup \Theta_{i2}(t))$, and $w_{ij}(t) = \delta(t)/|\Theta_{i1}(t) \cup \Theta_{i2}(t)|$ for $j \in \Theta_{i1}(t) \cup \Theta_{i2}(t)$, where $\delta(t)$ can be any decreasing sequence with limit being zero. This scenario is an approximation of taking arithmetic mean of neighbors used in many discrete-time consensus-based systems [2].

Remark 3. Our algorithms are purely distributed and have high flexibility as well as low complexity. The time complexity of the above strategies is dominated by the sorting task, which is effectively done by for instance Quicksort having complexity $O(n \ln n)$. Moreover, compared to the consensus seeking over bicolored graphs [22], [24], our systems (3) and (4) are not confined to be only continuous-time monotone systems.

III. Convergence analysis for continuous-time systems

Denote by $\underline{x}(t) = \min_{i \in C} x_i(t)$ the minimum state of cooperative nodes and $\overline{x}(t) = \max_{i \in C} x_i(t)$ the maximum state of cooperative nodes at time t. The following lemma says that $[\underline{x}(0), \overline{x}(0)]$ is a safety interval for all cooperative nodes under the proposed resilient consensus strategy.

Lemma 3. *Consider the system (1) over $G(V, E^o, E^p)$ following the resilient consensus strategy (3) with parameter r. If the number of Byzantine nodes is locally bounded by r, then $x_i(t) \in [\underline{x}(0), \overline{x}(0)]$ for all $i \in C$ and $t \geq 0$.*

Proof. We will only prove $x_i(t) \leq \overline{x}(0)$ for any $i \in C$ and $t \geq 0$. The other part of the lemma can be shown likewise. If the proposition is not true, then there must exist $i \in C$ and some time $\hat{t}$ such that (i) $x_j(t') \leq \overline{x}(0)$ for all $t' \leq \hat{t}$ and $j \in C$, and (ii) $x_i(\hat{t}) = \overline{x}(0)$ and $\dot{x}_i(\hat{t}) > 0$. In view of (3),

$$0 < \dot{x}_i(\hat{t}) = \sum_{j \in (N_i \cup \{i\}) \setminus R_i(\hat{t})} a_{ij} \rho_{ij}(\hat{t}) f_{ij}\left(x_j^i(\hat{t}), x_i(\hat{t})\right).$$

If $j \in C$, we have $x_i(\hat{t}) = \overline{x}(0) \geq x_j^i(\hat{t})$ by the above comment. If $j \in B$, the same estimate applies since node i has no more than r Byzantine neighbors and the algorithm ensures $x_j^i(\hat{t})$ is upper bounded by the value of a node in $((C \cap N_i) \cup \{i\}) \setminus R_i(\hat{t})$. Noting that $a_{ij} > 0$, $\rho_{ij}(\hat{t}) > 0$ and $f_{ij}\left(x_j^i(\hat{t}), x_i(\hat{t})\right) \leq 0$ by the assumptions of f_{ij}, the right-hand side of the above inequality is less than or equal to zero, which is a contradiction. Hence, we proved $x_i(t) \leq \overline{x}(0)$ for any $t \geq 0$. □

The network $G(V, E^o, E^p)$ is fixed. However, as nodes are adjusted as the system evolves according to the resilient consensus strategies, the network topology is essentially time-dependent. This can be seen in Equations (3) and (4), where $R_i(t)$ is a set dependent on time t. The following assumption captures the essential nature of time-varying topology due to dynamical censoring of values.

Assumption 1. Denote by $\{\tau_l\}_{l \geq 1}$ the ordered time sequence at which the set $R_i(t)$ alters for some $i \in C$. We assume there is $\tau > 0$ such that $\tau_{l+1} - \tau_l \geq \tau$ for all $l \geq 1$.

Theorem 1. *Consider the system (1) over $G(V, E^o, E^p)$ following the resilient consensus strategy (3) with parameter*

r. Suppose Assumption 1 holds and $G(V,E^o,E^p)$ is $(r+1)$-robust. If the number of Byzantine nodes is locally bounded by r, then resilient consensus can be achieved.

Proof. The upper Dini derivative of a continuous function f is given by

$$d^+f(t)=\limsup_{s\to 0+}\frac{f(t+s)-f(t)}{s}.$$

In light of the property of upper Dini derivative [37], the derivatives of $\overline{x}(t)$ and $\underline{x}(t)$ along the solution of (3) are described as

$$\begin{aligned}d^+\overline{x}(t)&=\dot{x}_{\bar{i}}(t)\\&=\sum_{j\in(N_{\bar{i}}\cup\{\bar{i}\})\backslash R_{\bar{i}}(t)}a_{\bar{i}j}\rho_{\bar{i}j}(t)f_{\bar{i}j}\left(x_j^{\bar{i}}(t),x_{\bar{i}}(t)\right)\end{aligned}\tag{5}$$

and

$$\begin{aligned}d^+\underline{x}(t)&=\dot{x}_{\underline{i}}(t)\\&=\sum_{j\in(N_{\underline{i}}\cup\{\underline{i}\})\backslash R_{\underline{i}}(t)}a_{\underline{i}j}\rho_{\underline{i}j}(t)f_{\underline{i}j}\left(x_j^{\underline{i}}(t),x_{\underline{i}}(t)\right),\end{aligned}\tag{6}$$

where $\dot{x}_{\bar{i}}(t):=\max_{i\in\overline{I}(t)}\dot{x}_i(t)$ and $\dot{x}_{\underline{i}}(t):=\max_{i\in\underline{I}(t)}\dot{x}_i(t)$, and the index sets are given by $\overline{I}(t)=\{i\in C:x_i(t)=\overline{x}(t)\}$ and $\underline{I}(t)=\{i\in C:x_i(t)=\underline{x}(t)\}$. Define the gap between the maximum and minimum of states as $\Delta(t):=\overline{x}(t)-\underline{x}(t)$. Consider the right-hand side of (5). If $j\in C$, by definition we obtain $x_j^{\bar{i}}(t)\le x_{\bar{i}}(t)$. If $j\in B$, we examine two cases: (i) $(j,\bar{i})\in E^p\backslash E^o$ and $x_j^{\bar{i}}(t)\ge x_{\bar{i}}(t)$ (ii) otherwise. For (i), $a_{\bar{i}j}\rho_{\bar{i}j}(t)f_{\bar{i}j}\left(x_j^{\bar{i}}(t),x_{\bar{i}}(t)\right)$ tends to zero. For (ii), $x_j^{\bar{i}}(t)$ is upper bounded by the value of a node in $((C\cap N_{\bar{i}})\cup\{\bar{i}\})\backslash R_{\bar{i}}(t)$ and hence upper bounded by $x_{\bar{i}}(t)$. Involving the assumptions of $\rho_{\bar{i}j}(t)$ and $f_{\bar{i}j}$, we know the right-hand side of (5) is at most zero when t is sufficiently large. By an analogous reasoning, the right-hand side of (6) is at least zero. Therefore,

$$d^+\Delta(t)=d^+\overline{x}(t)-d^+\underline{x}(t)\le 0$$

for sufficiently large t.

Next we claim that $d^+\Delta(t)$ tends to zero as t goes to infinity. We use a contradiction argument. There are two constants $\varepsilon<0$ and $\delta>0$ as well as an ordered time sequence $\{\sigma_l\}_{l\ge1}$ with $\lim_{l\to\infty}\sigma_l=\infty$ satisfying $d^+\Delta(\sigma_l)\le 2\varepsilon$ and $\sigma_{l+1}-\sigma_l>\delta$ for all $l\ge 1$. Consider an interval I satisfying $I\cap\{\tau_l\}_{l\ge1}=\emptyset$, where $\{\tau_l\}_{l\ge1}$ is defined in Assumption 1. As $\dot{x}_i(t)$ is bounded for any node $i\in C$ and $d^+\Delta(t)$ is continuous on I, $d^+\Delta(t)$ must be uniformly continuous on I. Hence, we can find $\hat{\delta}>0$ such that for $t_2>t_1$ in I with $t_2-t_1<\hat{\delta}$, we obtain $|d^+\Delta(t_1)-d^+\Delta(t_2)|<-\varepsilon$. In light of Assumption 1, we can choose $\hat{\delta}>0$ small enough so that for any $l>1$, we have $[\sigma_l-\hat{\delta},\sigma_l+\hat{\delta}]\subseteq I$ for some I defined as above. For $t\in[\sigma_l-\hat{\delta},\sigma_l+\hat{\delta}]$, the following estimate holds:

$$\begin{aligned}d^+\Delta(t)=&-|d^+\Delta(\sigma_l)-(d^+\Delta(\sigma_l)-d^+\Delta(t))|\\\le&-(|d^+\Delta(\sigma_l)|-|d^+\Delta(\sigma_l)-d^+\Delta(t)|)\\\le&2\varepsilon-\varepsilon=\varepsilon.\end{aligned}\tag{7}$$

Note that we can choose $\hat{\delta}$ sufficiently small so that $\{[\sigma_l-\hat{\delta},\sigma_l+\hat{\delta}]\}_{l\ge1}$ are mutually exclusive. It follows from the fact $d^+\Delta(t)\le 0$ for some sufficiently large $T>0$ and (7) that

$$\begin{aligned}\int_T^\infty d^+\Delta(t)dt\le&\lim_{\ell\to\infty}\sum_{l=1}^{\ell}\int_{\sigma_l-\hat{\delta}}^{\sigma_l+\hat{\delta}}d^+\Delta(t)dt\\\le&2\varepsilon\hat{\delta}\cdot\lim_{\ell\to\infty}\ell\\=&-\infty,\end{aligned}$$

which provides contradiction with $\Delta(t)\ge0$ for all $t\ge0$. We have proved $d^+\Delta(t)$ tends to zero as $t\to\infty$.

This together with the analysis built on (5) and (6) implies that there are two constants $\overline{y}\ge\underline{y}$ satisfying $\lim_{t\to\infty}\overline{x}(t)=\lim_{t\to\infty}x_{\bar{i}}(t)=\overline{y}$ and $\lim_{t\to\infty}\underline{x}(t)=\lim_{t\to\infty}x_{\underline{i}}(t)=\underline{y}$. Assume that $\overline{y}>\underline{y}$. Therefore, there exists time $\hat{t}>0$ so that for any $t>\hat{t}$, $x_{\bar{i}}(t)>\overline{y}-\delta>\underline{y}+\delta>x_{\underline{i}}(t)$ for some $\delta>0$. By assumption $G(V,E^o,E^p)$ is $(r+1)$-robust, invoking Lemma 1 and Lemma 2 we know that $G(V,E^o,E^p)$ is pseudo-strongly connected. It follows from (3) and $\lim_{t\to\infty}\dot{x}_{\bar{i}}(t)=0$ that $\lim_{t\to\infty}x_j^{\bar{i}}(t)-x_{\bar{i}}(t)=0$ for any $j\in(N_{\bar{i}}\cup\{\bar{i}\})\backslash R_{\bar{i}}(t)$. Analogously, $\lim_{t\to\infty}\dot{x}_{\underline{i}}(t)=0$ implies $\lim_{t\to\infty}x_j^{\underline{i}}(t)-x_{\underline{i}}(t)=0$ for any $j\in(N_{\underline{i}}\cup\{\underline{i}\})\backslash R_{\underline{i}}(t)$. As V contains only a finite number of nodes, there exists $t'>\hat{t}$ such that in the communication topology at t', there is a directed path from $k_{\bar{i}\underline{i}}$ to $\bar{i}$ and a directed path from $k_{\bar{i}\underline{i}}$ to $\underline{i}$, and $x_{k_{\bar{i}\underline{i}}}(t')<\underline{y}+\delta$ and $x_{k_{\bar{i}\underline{i}}}(t')>\overline{y}-\delta$. Clearly, this is a contradiction. Hence, we must have $\overline{y}=\underline{y}$. □

Remark 4. From the perspective of graph topology, previous works typically assumes that $G(V,E)$ is $(2r+1)$-robust in the presence of r-locally bounded Byzantine nodes; c.f. [8], [11], [12]. This condition is obviously stronger than our assumption of $(r+1)$-robustness.

The above proposed algorithm can be modified to deal with scaled consensus problems [11], [38] defined as below.

Definition 7 (resilient scaled consensus). Given $\gamma_i\ne0$ for $i\in V$. The cooperative agents in $G(V,E)$ are said to achieve resilient scaled consensus with respect to $(\gamma_1,\cdots,\gamma_n)$ if, for any initial configuration $\{x_i(0)\}_{i\in V}$,

$$\lim_{t\to\infty}\gamma_ix_i(t)-\gamma_jx_j(t)=0$$

for all $i,j\in C$.

When $\gamma_i\equiv1$ for all $i\in V$, we readily reproduce the consensus problem. In general, scaled consensus requires the ratio of cooperative agents converges to a pre-assigned (possibly negative) value, naturally delineating the cooperative and antagonistic interactions between individuals. For agents with continuous-time dynamics, the resilient consensus protocol proposed in Section II.A can be modified as follows.

Given the parameter r, at time t each cooperative agent $i\in C$ receives the state values $\{\gamma_jx_j^i(t)\}_{j\in N_i}$ from its neighbors and determines the E^o and E^p layers for its incoming edges. Sort respectively the two lists $\{\gamma_jx_j^i(t)\}_{j\in N_i^o}$ and $\{\gamma_jx_j^i(t)\}_{j\in N_i^p}$ in a descending manner. Remove up to r highest values in $\{\gamma_jx_j^i(t)\}_{j\in N_i^o}$ which are strictly larger than $\gamma_ix_i(t)$, and up to r lowest values in $\{\gamma_jx_j^i(t)\}_{j\in N_i^p}$ which are strictly smaller than $\gamma_ix_i(t)$. All removed indices are recorded in a set $R_i(t)$ (see the Removal Algorithm 2 given below). We instantiate f_i in (1) for $t \geq 0$ as

$$\dot{x}_i(t) = \operatorname{sgn}(\gamma_i) \sum_{j \in (N_i \cup \{i\}) \setminus R_i(t)} a_{ij} \rho_{ij}(t) f_{ij}\left(\gamma_j x_j^i(t), \gamma_i x_i(t)\right),$$

where the signum function $\operatorname{sgn}(b) = 1$ if $b > 0$ and $\operatorname{sgn}(b) = -1$ if $b < 0$, and ρ_{ij} and f_{ij} are as defined before. We can prove the following result with a similar proof as in Theorem 1.

Removal Algorithm 2 for a cooperative agent $i \in C$

Input: $\gamma_i x_i(t)$, $\{\gamma_j x_j^i(t)\}_{j \in N_i^o}$, $\{\gamma_j x_j^i(t)\}_{j \in N_i^p}$
Output: $R_i(t)$
sort $\{\gamma_j x_j^i(t)\}_{j \in N_i^o}$ decreasingly as
  $L^o := \{\gamma_{j_1} x_{j_1}^i(t), \gamma_{j_2} x_{j_2}^i(t), \cdots, \gamma_{j_{|N_i^o|}} x_{j_{|N_i^o|}}^i(t)\}$
let $R_i(t) = \emptyset$ and
  $\bar{\ell} = \max\{1 \leq l \leq |N_i^o| : \gamma_{j_l} x_{j_l}^i(t) > \gamma_i x_i(t)\}$
if $\bar{\ell} > r$
  $\bar{\ell} = r$
end if
for $l = 1$ till $l = \bar{\ell}$
  add v_{j_l} into $R_i(t)$
end for
sort $\{\gamma_j x_j^i(t)\}_{j \in N_i^p}$ decreasingly as
  $L^p := \{\gamma_{j_1} x_{j_1}^i(t), \gamma_{j_2} x_{j_2}^i(t), \cdots, \gamma_{j_{|N_i^p|}} x_{j_{|N_i^p|}}^i(t)\}$
let $\underline{\ell} = \min\{1 \leq l \leq |N_i^p| : \gamma_{j_l} x_{j_l}^i(t) < \gamma_i x_i(t)\}$
if $|N_i^p| - \underline{\ell} + 1 > r$
  $\underline{\ell} = |N_i^p| - r + 1$
end if
for $l = \underline{\ell}$ till $l = |N_i^p|$
  add v_{j_l} into $R_i(t)$
end for

Corollary 1. *Consider the system (1) over $G(V, E^o, E^p)$ following the above resilient scaled consensus strategy with parameter r. Suppose Assumption 1 holds and $G(V, E^o, E^p)$ is $(r+1)$-robust. If the number of Byzantine nodes is locally bounded by r, then resilient scaled consensus with respect to $(\gamma_1, \cdots, \gamma_n)$ can be achieved.*

IV. Convergence analysis for discrete-time systems

Recall that $\underline{x}(t) = \min_{i \in C} x_i(t)$ and $\overline{x}(t) = \max_{i \in C} x_i(t)$ are, respectively, the minimum state and maximum state of cooperative nodes at t. The following lemma indicates that $\{[\underline{x}(t), \overline{x}(t)]\}_{t \geq 1}$ is a nested sequence with respect to time under the proposed resilient consensus strategy.

Lemma 4. *Consider the system (2) over $G(V, E^o, E^p)$ following the resilient consensus strategy (4) with parameter r. If the number of Byzantine nodes is locally bounded by r, then $x_i(t+1) \in [\underline{x}(t), \overline{x}(t)]$ for all $i \in C$ and $t \geq 0$.*

Proof. The proposed resilient consensus strategy implies that $x_i(t+1)$ is a convex combination of values $x_j^i(t)$ with $j \in (N_i \cup \{i\}) \setminus R_i(t)$. In view of the locally bounded Byzantine nodes and the choice of $w_{ij}(t)$, it is clear that all non-zero constitutes of the convex combination in (4) are within the interval $[\underline{x}(t), \overline{x}(t)]$. Hence, $x_i(t+1) \in [\underline{x}(t), \overline{x}(t)]$. □

Theorem 2. *Consider the system (2) over $G(V, E^o, E^p)$ following the resilient consensus strategy (4) with parameter r. Suppose $G(V, E^o, E^p)$ is $(r+1)$-robust. If the number of Byzantine nodes is locally bounded by r, then resilient consensus can be achieved.*

Proof. Thanks to Lemma 4, we can define $\overline{y} := \lim_{t \to \infty} \overline{x}(t)$ and $\underline{y} := \lim_{t \to \infty} \underline{x}(t)$. Obviously, $\overline{y} \geq \underline{y}$. We will show the equality holds by method of contradiction. Suppose $\overline{y} > \underline{y}$. We choose $\delta_0 > 0$ with $\overline{y} - \delta_0/2 > \underline{y} + \delta_0/2$. We consider two sequences of subsets of cooperative nodes $\overline{B}(t, \delta_l) := \{i \in C : x_i(t) > \overline{y} - \delta_l/2\}$ and $\underline{B}(t, \delta_l) := \{i \in C : x_i(t) < \underline{y} + \delta_l/2\}$ for given $t, \delta_l > 0$ with $l \geq 0$. Note that $\overline{B}(t, \delta_0)$ and $\underline{B}(t, \delta_0)$ are mutually exclusive by our definition. We choose $\delta \in (0, \delta_0/2)$ such that $\delta < w^{n-2r}\delta_0/(2(1 - w^{n-2r}))$, where $w > 0$ defined in Section II.A can be taken as $w \in (0, 1)$. For this δ, we take time $t_\delta > 0$ such that $\overline{x}(t) < \overline{y} + \delta$ and $\underline{x}(t) > \overline{y} - \delta$ for any $t \geq t_\delta$ by using the convergence. Let $\hat{G}(t_\delta) := G(V, E^o(t_\delta), E^p(t_\delta))$ be the communication topology at time t_δ, which according to our algorithm is obtained by each cooperative node removing up to r incoming edges in E^o as well as removing up to r incoming edges in E^p. By Lemma 1, $\hat{G}(t_\delta)$ is 1-robust.

Consider the two nonempty mutually exclusive sets $\overline{B}(t_\delta, \delta_0)$ and $\underline{B}(t_\delta, \delta_0)$. There exists a cooperative agent in $\overline{B}(t_\delta, \delta_0)$ or in $\underline{B}(t_\delta, \delta_0)$, which has at least one neighbor in $E^o(t_\delta)$ or in $E^p(t_\delta)$ outside its set. Without loss of generality, we assume $i \in \overline{B}(t_\delta, \delta_0)$. If i has one neighbor in $E^p(t_\delta)$ outside its set $\overline{B}(t_\delta, \delta_0)$, the value of this neighbor is upper bounded by $\overline{y} - \delta_0/2$ and it is used by node i in the update. [Notice that i has no neighbor in $E^o(t_\delta)$ outside its set. Otherwise, by the definition of $\hat{G}(t_\delta)$ this value will be used by node i, but on the other hand this value cannot be used by node i due to (4), which leads to a contradiction.] Therefore, using the resilient consensus strategy we derive the following estimate

$$\begin{aligned} x_i(t_\delta + 1) \leq& (1 - w)\overline{x}(t_\delta) + w(\overline{y} - \delta_0/2) \\ \leq& (1 - w)(\overline{y} + \delta) + w(\overline{y} - \delta_0/2) \\ =& \overline{y} - \delta_0 w/2 + (1 - w)\delta. \end{aligned} \tag{8}$$

It is easily see that (8) still applies to any cooperative agent outside $\overline{B}(t_\delta, \delta_0)$ since such nodes will leverage their own value in (4), which is again upper bounded by $\overline{y} - \delta_0/2$. In a similar manner, if $i \in \underline{B}(t_\delta, \delta_0)$, we can obtain the estimate

$$x_i(t_\delta + 1) \geq \underline{y} + \delta_0 w/2 - (1 - w)\delta. \tag{9}$$

This is also satisfied by cooperative agents outside $\underline{B}(t_\delta, \delta_0)$.

We take $\delta_1 = w\delta_0 - 2(1 - w)\delta$ and $0 < \delta < \delta_1/2 < \delta_0/2$ holds. Since the two sets $\overline{B}(t_\delta + 1, \delta_1)$ and $\underline{B}(t_\delta + 1, \delta_1)$ are mutually exclusive (by (8) and (9)), we arrive at $|\overline{B}(t_\delta + 1, \delta_1)| < |\overline{B}(t_\delta, \delta_0)|$ or $|\underline{B}(t_\delta + 1, \delta_1)| < |\underline{B}(t_\delta, \delta_0)|$ based on the above arguments. For $l \geq 1$, set $\delta_l = w\delta_{l-1} - 2(1 - w)\delta$ and we have $\delta_{l-1} > \delta_l$. Repeatedly applying the above comments, we know that there must exist some $T \leq n - 2r$ such that $\overline{B}(t_\delta + T, \delta_T)$ or $\underline{B}(t_\delta + T, \delta_T)$ is an empty set. This is because

$\hat{G}(t_\delta + T)$ has at most $n - 2r$ nodes. Using our definition of δ, we estimate

$$\begin{aligned}\delta_T =& w\delta_{T-1} - 2(1-w)\delta \\ =& w^T\delta_0 - 2(1-w)(1 + w + \cdots + w^{T-1})\delta \\ =& w^T\delta_0 - 2(1-w^T)\delta \\ \geq& w^{n-2r}\delta_0 - 2(1-w^{n-2r})\delta \\ >& 0.\end{aligned}$$

Therefore, either all cooperative agents at time $t_\delta + T$ have values no more than $\overline{y} - \delta_T/2 < \overline{y}$ or all of them have values no less than $\underline{y} + \delta_T/2 > \underline{y}$. This contradicts the definition of $\overline{y}$ or $\underline{y}$. Hence, we must have $\overline{y} = \underline{y}$. □

Given $\gamma_i \neq 0$ for $i \in V$, we modify the resilient consensus strategy for discrete-time dynamical agents similarly as in Section III, and we instantiate f_i in (2) for $t \geq 0$ as

$$x_i(t+1) = \mathrm{sgn}(\gamma_i) \sum_{j \in (N_i \cup \{i\}) \setminus R_i(t)} w_{ij}(t)\gamma_j x_j^i(t),$$

where $\mathrm{sgn}(\cdot)$ is the signum function, the weights $w_{ij}(t)$ are defined as before. Here, we assume $\sum_{j \in (N_i \cup \{i\}) \setminus R_i(t)} |\gamma_i| w_{ij}(t) = 1$.

By a similar proof as in Theorem 2, we can show the following result for resilient scaled consensus.

Corollary 2. *Consider the system (2) over $G(V, E^o, E^p)$ following the above resilient scaled consensus strategy with parameter r. Suppose $G(V, E^o, E^p)$ is $(r+1)$-robust. If the number of Byzantine nodes is locally bounded by r, then resilient scaled consensus with respect to $(\gamma_1, \cdots, \gamma_n)$ can be achieved.*

V. Numerical Examples

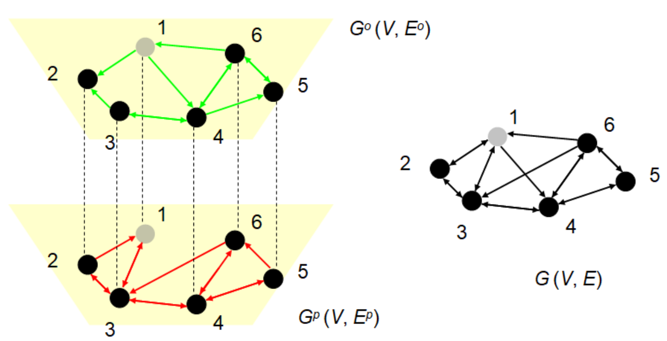

Fig. 1. Left: A multiplex network $G(V, E^o, E^p)$ with $B = \{1\}$ and $C = \{2, 3, \cdots, 6\}$ for Example 1. Green and red arrows represent optimistic and pessimistic edges respectively. Right: The one-mode version $G(V, E)$ with $E = E^o \cup E^p$.

Example 1. Here, we consider a multiplex network $G(V, E^o, E^p)$ over $V = \{1, 2, \cdots, 6\}$, where 1 is a Byzantine node and the other nodes are cooperative; see Fig. 1 for an illustration. The one-mode version $G(V, E)$ has a binary adjacency matrix A. To combat the malicious agent, we are interested in the implementation of our resilient consensus strategies having parameter $r + 1$ with $r = 1$. It is direct to check that $G(V, E^o, E^p)$ is 2-robust. However, since for example $N_5 = \{4, 6\}$ and $N_6 = \{4, 5\}$, $G(V, E)$ is not 3-robust. As a result, neither $G^o(V, E^o)$ nor $G^p(V, E^p)$ is 3-robust. Therefore, we cannot conclude that resilient consensus will be reached over $G(V, E)$ with any previously known consensus results. In fact, these networks are sparser than the previous requirement over robustness, namely, $(2r+1)$-robustness; c.f. [6], [8]–[12].

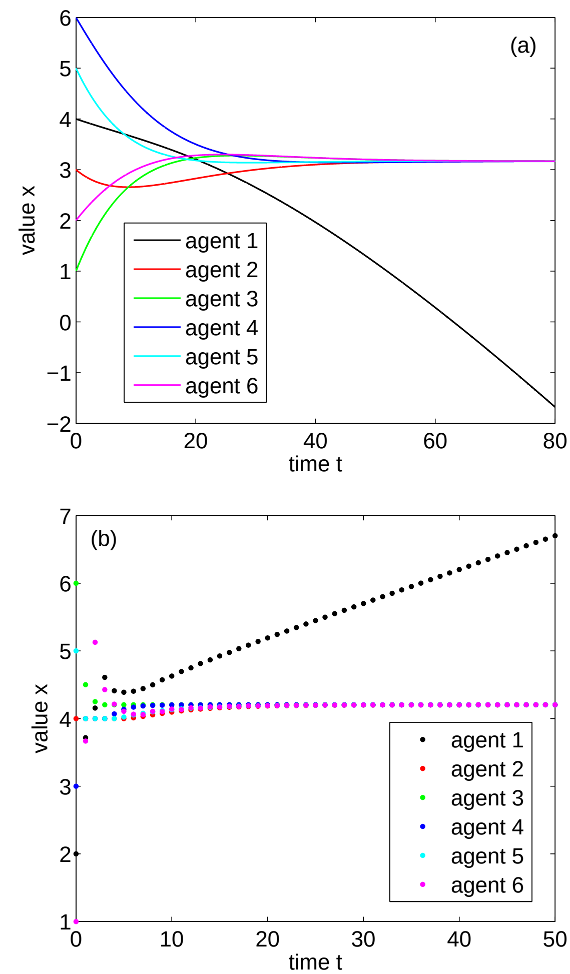

Fig. 2. Resilient consensus over directed network $G(V, E^o, E^p)$ of Example 1 in the presence of Byzantine node 1 for (a) continuous-time multi-agent system (1) with $\rho_{ij}(t) = t^{-2}$ and (b) discrete-time multi-agent system (2) with $\delta(t) = t^{-1}$.

In Fig. 2(a) we show the state evolution of the multi-agent system having continuous-time agents with initial conditions $x_1(0) = 4$, $x_2(0) = 3$, $x_3(0) = 1$, $x_4(0) = 6$, $x_5(0) = 5$, $x_6(0) = 2$. The Byzantine node 1 follows its own dynamics $\dot{x}_1(t) = (x_6(t) - x_1(t))/50 - t/400$ while cooperative nodes follow the strategy described in Remark 2 with $f_{ij}(b_i, b_j) = 0.1 \cdot (b_i - b_j)$ and $\rho_{ij}(t) = \alpha t^{-2}$ with $\alpha = 1$ for appropriate nodes. We observe from Fig. 2(a) that all cooperative agents are able to achieve consensus as one would expect from Theorem 1.

The trajectories of discrete-time dynamical agents with initial conditions $x_1(0) = 2, x_2(0) = 4, x_3(0) = 6, x_4(0) = 3$, $x_5(0) = 5$, $x_6(0) = 1$ are shown in Fig. 2(b). The Byzantine node 1 follows the update rule $x_1(t+1) = (x_2(t) + x_3(t) +$

$x_6(t))/3 + t/20$ and cooperative nodes follow the adaptive strategy described in Remark 2 with $\delta(t) = \beta t^{-1}$ and $\beta = 1$. The consensus result shown in Fig. 2(b) is consistent with the theoretical prediction of Theorem 2.

	$\alpha = 1$	$\alpha = 5$	$\alpha = 10$	$\alpha = t$
$t^*_{0.0001}$	71.36	73.01	75.22	106.95
	$\beta = 1$	$\beta = 2$	$\beta = 3$	$\beta = \sqrt{t}$
$t^*_{0.0001}$	28	28	29	34

TABLE I
CONSENSUS TIME $t^*_{0.001}$ FOR EXAMPLE 1 WITH $\rho_{ij}(t) = \alpha t^{-2}$ IN CONTINUOUS-TIME SYSTEM (1) AND $\delta(t) = \beta t^{-1}$ IN DISCRETE-TIME SYSTEM (2).

Next, we consider the influence of choices of α and β on the convergence speed of the proposed protocols. To this end, we define the consensus time for the multi-agent system (1) or (2) as

$$t^*_\varepsilon := \min\left\{t \geq 0 : \max_{i,j\in C} |x_i(t) - x_j(t)| \leq \varepsilon\right\} \tag{10}$$

for a small $\varepsilon > 0$. In Table 1 we summarize the consensus time t^*_ε with $\varepsilon = 10^{-3}$ for different values of α and β in both the continuous-time and discrete-time protocols as described as above. The consensus time is seen to be increasing with respect to α and β in general. This is consistent with the analysis in our main results, where the convergence of ρ_{ij} and δ is essential in containing the unwanted perturbation.

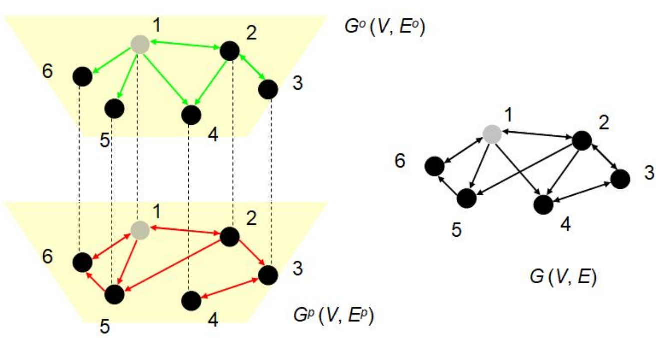

Fig. 3. Left: A multiplex network $G(V, E^o, E^p)$ with $B = \{1\}$ and $C = \{2, 3, \cdots, 6\}$ for Example 2. Green and red arrows represent optimistic and pessimistic edges respectively. Right: The one-mode version $G(V, E)$ with $E = E^o \cup E^p$.

Example 2. In the following, we consider a sparser multiplex network $G(V, E^o, E^p)$ over $V = \{1, 2, \cdots, 6\}$ (see Fig. 3), where 1 is a Byzantine node and the other nodes are cooperative. We assume a binary adjacency matrix A for the one-mode version $G(V, E)$. As before, it is straightforward to check that $G(V, E^o, E^p)$ is 2-robust while $G(V, E)$ is not 3-robust. This is because $N_6 = \{1, 5\}$ and hence the two disjoint subsets $\{1, 5\}$ and $\{2, 3, 4, 6\}$ do not satisfy the 3-robustness condition in [8]. Moreover, with slightly more effort we know that $G(V, E)$ even fails to be 2-robust (for example by scoping the two subsets $\{1, 6\}$ and $\{3, 4\}$). Therefore, this network satisfies the assumptions in our Theorems 1 and 2 with the parameter $r = 1$, but it does not meet the requirement of the typical fault-tolerant filtering strategies with $r = 1$ Byzantine node in [8].

We take initial conditions $x_1(0) = 4$, $x_2(0) = 3$, $x_3(0) = 1$, $x_4(0) = 6$, $x_5(0) = 5$, $x_6(0) = 2$ as in Example 1. For continuous-time multi-agent system, we make Byzantine node 1 follows its own dynamics $\dot{x}_1(t) = (x_6(t) - x_1(t))/50 - t/400$ while cooperative nodes follow the strategy described in Remark 2 with $f_{ij}(b_i, b_j) = 0.1 \cdot (b_i - b_j)$ and $\rho_{ij}(t) = t^{-2}$ for appropriate nodes. The state evolution is shown in Fig. 4(a). For discrete-time multi-agent system, we take $x_1(0) = 2$, $x_2(0) = 4$, $x_3(0) = 6$, $x_4(0) = 3$, $x_5(0) = 5$, $x_6(0) = 1$ as in Example 1 and assume that the Byzantine node 1 follows the update rule $x_1(t + 1) = (x_2(t) + x_6(t))/2 + t/20$ and cooperative nodes follow the adaptive strategy described in Remark 2 with $\delta(t) = t^{-1}$. The result is shown in Fig. 4(b).

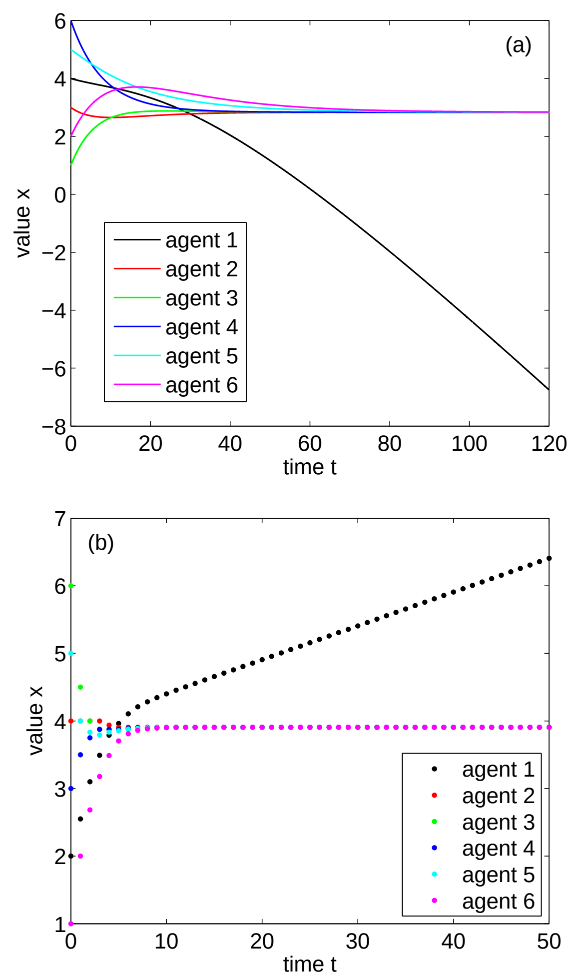

Fig. 4. Resilient consensus over directed network $G(V, E^o, E^p)$ of Example 2 in the presence of Byzantine node 1 for (a) continuous-time multi-agent system (1) with $\rho_{ij}(t) = t^{-2}$ and (b) discrete-time multi-agent system (2) with $\delta(t) = t^{-1}$.

We observe that the cooperative nodes 2-6 are able to reach consensus in spite of the presence of the Byzantine node 1 as predicted by Theorem 1 and Theorem 2, respectively. The network $G(V, E)$ here has 13 directed edges, which is obviously sparser than that in Example 1 containing 17 directed edges. Comparing Fig. 4(a) with Fig. 2(a), and Fig. 4(b) with

Fig. 2(b), we draw two interesting observations. Firstly, denser communication graph in one-mode (i.e., aggregated) version does not necessarily lead to faster convergence speed. The rate of reaching consensus depends in general on the specific connection architecture in the separate layers of the multiplex network. For instance, in Fig. 4(a) the system converges at around $t = 100$ while it converges at around $t = 70$ in Fig. 2(a); in Fig. 4(b) the convergence time is at around $t = 15$ while it is at around $t = 25$ in Fig. 2(b). We have done more simulations with different initial conditions and dynamics, which confirm this non-monotonicity in general. Secondly, when we compare the final consensus values as well as the trajectories (for both cooperative and Byzantine agents) in Fig. 2 and Fig. 4, we notice that they are different in both continuous-time and discrete-time cases. This phenomenon reveals that the network structure mutation, under our resilient consensus strategies, has an impact not only on the transient trajectories of the agents but also on their ultimate equilibrium state.

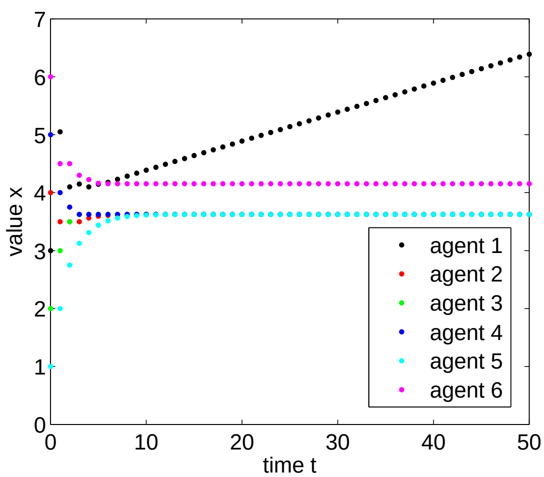

Fig. 5. State evolution over directed network $G(V, E)$ in Fig. 3 in the presence of Byzantine node 1 with the resilient consensus strategy given in [9, Theorem 2].

Finally, we consider a classical resilient consensus algorithm proposed in the work [9] for discrete-time agent dynamics with the same Byzantine node behavior as before. The underlying communication network is taken as the one-mode graph in Fig. 3. The system state evolution is shown in Fig. 5. According to the discussion in the beginning of Example 2, $G(V, E)$ is not 2-robust. Therefore, Theorem 2 in [9] is no longer a sufficient condition for resilient consensus in this case. Our Fig. 5 depicts such a counter example. This again highlights the usefulness of our multiplex network framework, which allows resilient consensus in a weaker connectivity condition.

VI. Concluding remarks

In this paper, we have considered consensus problems in the presence of locally bounded Byzantine nodes featuring asymmetric confidence intervals. We introduce a novel multiplex network presentation and associated concepts of robustness and pseudo-strongly connectivity. Adaptive distributed resilient protocols are designed and applied to both continuous-time and discrete-time multi-agent systems to ensure consensus among cooperative agents over directed networks. Capitalizing on the multiplex network framework, we are able to establish much milder robustness conditions than the existing works. The results have been extended to accommodate resilient scaled consensus problems, where both cooperative consensus and antagonistic consensus are permitted. Numerical examples have been worked out to confirm our theoretical findings and demonstrate the flexibility of the multiplex consensus framework. An interesting future direction is to consider vector states [39], [40], which offer an opportunity to generate the current multiplex network framework from two layers to more layers.

To accommodate the phenomenon of asymmetric confidence intervals, we in this paper scopes a solution by using multiplex networks. Another possible framework would be to consider a time-varying simplex network with the two sub-networks G^o and G^p periodically appearing. If at some time t an edge, for $(j, i) \in E^o \backslash E^p$ ($i \in C$) appears, the node j will be put into $R_i(t)$ whenever $x_j^i(t) < x_i(t)$. Compared to the ordinary W-MSR algorithms, more edges will be sieved in general. Potential fix include considering a stronger connectivity condition, restricting the occurrence of edges in $E^o \backslash E^p$ and $E^p \backslash E^o$, allowing certain fraction of trusted edges within N_i for all $i \in C$.

Finally, we mention that in many realistic applications, there are some other important aspects (such as delay, finite-time convergence, limited data rate [41], or packet dropout [42]) should be taken into consideration when designing resilient consensus algorithms. A very recent effort in finite-time resilient consensus is made by [43], where discontinuous system theory is applied to reach consensus in the symmetric confidence scenario.

Acknowledgment

The author thanks the handling editor and the anonymous reviewers for their constructive comments that helped improve the paper.

References

[1] R. Olfati-Saber, J. A. Fax, R. M. Murray, Consensus and cooperation in networked multi-agent systems. *Proc. IEEE*, 95(2007) 215–233.

[2] A. Jadbabaie, J. Lin, A. S. Morse, Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Trans. Autom. Contr.*, 48(2003) 988–1001.

[3] M. Mesbahi, M. Egerstedt, *Graph Theoretic Methods in Multiagent Networks*, Princeton University Press, New Jersey, 2010.

[4] J. Qin, Q. Ma, Y. Shi, L. Wang, Recent advances in consensus of multi-agent systems: a brief survey. *IEEE Trans. Ind. Electron.*, 64(2017) 4972–4983.

[5] H. Yang, Q.-L. Han, X. Ge, L. Ding, Y. Xu, B. Jiang, D. zhou, Fault-tolerant cooperative control of multiagent systems: a survey of trends and methodologies. *IEEE Trans. Ind. Inf.*, 16(2020) 4–17.

[6] Y. Shang, Hybrid consensus for averager-copier-voter networks with non-rational agents. *Chaos, Solitons & Fractals*, 110(2018) 244–251.

[7] D. Mukherjee, D. Zelazo, Robustness of consensus over weighted digraphs. *IEEE Trans. Netw. Sci. Eng.*, 6(2019) 657–670.

[8] H. J. LeBlanc, H. Zhang, X. Koutsoukos, S. Sundaram, Resilient asymptotic consensus in robust networks. *IEEE J. Select. Areas Commun.*, 31(2013) 766–781.

[9] H. Zhang, E. Fata, S. Sundaram, A notion of robustness in complex networks. *IEEE Trans. Contr. Netw. Syst.*, 2(2015) 310–320.

[10] S. M. Dibaji, H. Ishii, Resilient consensus of second-order agent networks: asynchronous update rules with delays. *Automatica*, 81(2017) 123–132.

[11] Y. Shang, Resilient consensus of switched multi-agent systems. *Syst. Contr. Lett.*, 122(2018) 12–18.

[12] Y. Shang, Consensus of hybrid multi-agent systems with malicious nodes. *IEEE Trans. Circuits Syst. Express Briefs*, 67(2020) 685–689.

[13] J. Huang, Y. Wu, L. Chang, M. Tao, X. He, Resilient consensus with switching networks and heterogeneous agents. *Neurocomput.*, 341(2019) 70–79.

[14] S. M. Dibaji, H. Ishii, R. Tempo, Resilient randomized quantized consensus. *IEEE Trans. Autom. Contr.*, 63(2018) 2508–2522.

[15] Y. Shang, Resilient interval consensus in robust networks. *Int. J. Robust Nonlin. Contr.*, doi: 10.1002/rnc.5153.

[16] Y. Shang, Resilient cluster consensus of multiagent systems. *IEEE Trans. Syst. Man Cybern. Syst.*, doi: 10.1109/TSMC.2020.2997855.

[17] A. Mitra, S. Sundaram, Byzantine-resilient distributed observers for LTI systems. *Automatica*, 108(2019) 108487.

[18] W. Abbas, A. Laszka, X. Koutsoukos, Improving network connectivity and robustness using trusted nodes with application to resilient consensus. *IEEE Trans. Contr. Netw. Syst.*, 5(2018) 2036–2048.

[19] Y. Wu, X. He, Finite-time consensus-based clock synchronization under deception attacks. *IEEE Access*, 8(2020) 110748.

[20] F. Ghawash, W. Abbas, Leveraging diversity for achieving resilient consensus in sparse networks. *IFAC PapersOnLine*, 52-20(2019) 339–344.

[21] Y. Shang, Deffuant model with general opinion distributions: First impression and critical confidence bound. *Complexity*, 19(2013) 38–49.

[22] D. Angeli, S. Manfredi, Criteria for asymptotic clustering of opinion dynamics towards bimodal consensus. *Automatica*, 103(2019) 230–238.

[23] A. G. Dimakis, S. Kar, J. M. F. Moura, M. G. Rabbat, A. Scaglione, Gossip algorithms for distriuted signal processing. *Proc. IEEE*, 98(2010) 1847–1864.

[24] S. Manfredi, D. Angeli, Necessary and sufficient conditions for consensus in nonlinear monotone networks with unilateral interactions. *Automatica*, 77(2017) 51–60.

[25] C. Antonopoulos, Y. Shang, Opinion formation in multiplex networks with general initial distributions. *Sci. Rep.*, 8(2018) 2852.

[26] A. Aleta, Y. Moreno, Multilayer networks in a nutshell. *Ann. Rev. Cond. Matt. Phys.*, 10(2019), 45–62.

[27] P. Wang, G. Wen, X. Yu, W. Yu, T. Huang, Synchronization of multi-layer networks: from node-to-node synchronization to complete synchronization. *IEEE Trans. Circuit Syst. I: Reg. Pap.*, 66(2019) 1141–1152.

[28] I. Aldasoro, I. Alves, Multiplex interbank networks and systemic importance: An application to European data. *J. Finan. Stab.*, 35(2018) 17–37.

[29] Y. Zhuang, O. Yağan, Multistage complex contagions in random multiplex networks. *IEEE Trans. Contr. Netw. Syst.*, 7(2020) 410–421.

[30] N. Castro, C. S. Q. Siew, Contributions of modern network science to the cognitive sciences: revisiting research spirals of representation and process. *Proc. R. Soc. A*, 476(2020) 20190825.

[31] G. Mangioni, G. Jurman, M. D. Domenico, Multilayer flows in molecular networks identify biological modules in the human proteome. *IEEE Trans. Netw. Sci. Eng.*, 7(2020) 411–420.

[32] Y. Huang, H. Krim, A. Panahi, L. Dai, Community detection and improved detectability in multiplex networks. *IEEE Trans. Netw. Sci. Eng.*, 7(2020) 1697–1709.

[33] Y. Shang, Generalized k-core percolation on correlated and uncorrelated multiplex networks. *Phys. Rev. E*, 101(2020) 042306.

[34] E. Estrada, Communicability geometry of multiplexes. *New J. Phys.*, 21(2019) 015004.

[35] M. Kivelä, M. A. Porter, Isomorphisms in multilayer networks. *IEEE Trans. Netw. Sci. Eng.*, 5(2018) 198–211.

[36] F. Sorrentino, L. M. Pecora, L. Trajkovic, Group consensus in multilayer networks. *IEEE Trans. Netw. Sci. Eng.*, 7(2020) 2016–2026.

[37] J. M. Danskin, The theory of max-min, with applications. *SIAM J. Appl. Math.*, 14(1966) 641–664.

[38] S. Roy, Scaled consensus. *Automatica*, 51(2015) 259–262.

[39] H. Mendes, M. Herlihy, N. Vaidya, V. Garg, Multidimensional agreement in Byzantine systems. *Distrib. Comput.*, 28(2015) 423–441.

[40] X. Wang, S. Mou, S. Sundaram, A resilient convex combination for consensus-based distributed algorithms. *Numer. Alg. Contr. Optim.*, 9(2019) 269–281.

[41] M. Komareji, Y. Shang, R. Bouffanais, Consensus in topologically interacting swarms under communication constraints and time-delays. *Nonlin. Dyn.*, 93(2018) 1287–1300.

[42] X. Zhao, X. Zheng, C. Ma, R. Li, Distributed consensus of multiple Euler-Lagrange systems networked by sample-data information with transmission delays and data packet dropouts. *IEEE Trans. Autom. Sci. Eng.*, 14(2017) 1440–1450.

[43] J. Usevitch, D. Panagou, Resilient finite-time consensus: a discontinuous systems perspective. *Proc. Amer. Contr. Conf.*, Denver, CO, 2020, pp. 3285–3290.