Detection and Isolation of Link Failures under the Agreement Protocol

M. Amin Rahimian, Victor M. Preciado

Abstract—In this paper a property of the multi-agent consensus dynamics that relates the failure of links in the network to jump discontinuities in the derivatives of the output responses of the nodes is derived and verified analytically. At the next step, an algorithm for sensor placement is proposed, which would enable the designer to detect and isolate any link failures across the network based on the observed jump discontinuities in the derivatives of the responses of a subset of nodes. These results are explained through elaborate examples.

I. INTRODUCTION

Multi-agent network systems, which consist of a group of interacting dynamic agents, have found promising applications in areas such as motion coordination of robots [1]. Such cooperative dynamics over a network may be strongly affected by the network failures and this has motivated the study of network dynamics following the removal of some links or nodes [2], [3]. By and large, the study of failures is an important topic in network science and it has various practical implications [4]. Consensus or agreement protocol has been extensively investigated in the recent literature as a fundamental evolution law for multi-agent networks [5]. The papers [6] and [7] address the issue of detectability for single and multiple link failures in a multi-agent system under the agreement protocol, where it is pointed out that link failures are detectable for a class of directed graphs with rooted out-branchings. Additional conditions in terms of the inter-nodal distances to the observation points are also provided for the detectability of links. The chief aim of this paper is to provide a method for detection and isolation of single link failures in a network that evolves according to the linear agreement protocol, based on the output responses of a subset of nodes.

The remainder of this paper is organized as follows. Section II gives some preliminaries on sets and graph theory, and introduces the notation that is used throughout the paper. The main theorem that forms the analytic basis for the proposed detection method is stated and proved in Section III. Next in Section IV a set of algorithms are proposed for the effective selection of observation points in the network. These algorithms together with the theorem in Section III, enable the network designer to detect and isolate single link failures based on the observed jump discontinuities in the derivatives of the output responses of a subset of nodes. Illustrative examples and discussions in Section V elucidate the results and Section VI concludes the paper.

* The authors are with the Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104-6228 USA. (email: preciado@seas.upenn.edu).

II. SETS AND DIGRAPHS

Throughout the paper, $\emptyset$ is the empty set, $\mathbb{N}$ denotes the set of all natural numbers, $\mathbb{R} = \mathbb{R} \cup \{\}$, and $\mathbb{R}$ denotes the set of all real numbers. Also, the set of integers $\{1, 2, \ldots, k\}$ is denoted by $\mathbb{N}_k$, and any other set is represented by a curved capital letter. The cardinality of a set $\mathcal{X}$, which is the number of its elements, is denoted by $|\mathcal{X}|$, and $\mathcal{P}(\mathcal{X}) = \{\mathcal{M}; \mathcal{M} \subset \mathcal{X}\}$ denotes the power-set of $\mathcal{X}$, which is the set of all its subsets. The difference of two sets $\mathcal{X}$ and $\mathcal{Y}$ is denoted by $\mathcal{X} \setminus \mathcal{Y}$ and is defined as $\{x; x \in \mathcal{X} \land x \notin \mathcal{Y}\}$, where $\land$ is the logical conjunction. In addition the logical implication and bi-implication are denoted by $\rightarrow$ and $\leftrightarrow$, respectively. Matrices are represented by capital letters, vectors are expressed by boldface lower-case letters, and the superscript $T$ denotes the matrix transpose. Moreover, $\mathcal{I}$ identifies the identity matrix with proper dimension, and the determinant of a matrix $\mathcal{D}$ is denoted by $\det(\mathcal{D})$, while $|\mathcal{D}|_{ij}$ indicates the element of $\mathcal{D}$ which is located at its $i-$th row and $j-$th column.

A directed graph or digraph is defined as an ordered pair of sets $\mathcal{G} := (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{\nu_1, \ldots, \nu_n\}$ is a set of $n = |\mathcal{V}|$ vertices and $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ is a set of directed edges. In the graphical representations, each edge $\epsilon := (\tau, \nu) \in \mathcal{E}$ is depicted by a directed arc from vertex $\tau \in \mathcal{V}$ to vertex $\nu \in \mathcal{V}$. Vertices $\nu$ and $\tau$ are referred to as the head and tail of the edge $\epsilon$, respectively; and if $\tau = \nu$, then $\epsilon$ is dubbed a self-loop. Given a set of vertices $\mathcal{X} \subset \mathcal{V}$, the set of all edges for which the heads belong to $\mathcal{X}$ but the tails do not, is referred to as the in-cut of $\mathcal{X}$, and is denoted by $\partial^{\mathcal{G}}_{\mathcal{X}} \mathcal{X} \subset \mathcal{E}$. The cardinality of $\partial^{\mathcal{G}}_{\mathcal{X}} \mathcal{X}$ is called the in-degree of $\mathcal{X}$, and is characterized as $\partial^{\mathcal{G}}_{\mathcal{X}} \mathcal{X} = |\partial^{\mathcal{G}}_{\mathcal{X}} \mathcal{X}|$. Notice that by definition there are no parallel arcs in the graphical representation described above. In other words, if two edges share the same pair of head and tail, then they are identical.

A matrix $\mathcal{W} \in \mathbb{R}^{(|\mathcal{V}| \times |\mathcal{V}|)}$ is called an in-weighting on $\mathcal{G}$ if $\forall \{\nu_i, \nu_j\} \subset \mathcal{V}, (\nu_j, \nu_i) \notin \mathcal{E} \rightarrow [\mathcal{W}]_{ij} = 0$. For a given digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and any pair of vertices $\{\nu_i, \nu_j\} \subset \mathcal{V}$, let edge $\epsilon := (\nu_j, \nu_i) \in \mathcal{E} \cup \{\epsilon\}$. The edge-index of $\epsilon$ is defined as a $|\mathcal{V}| \times |\mathcal{V}|$ matrix with exactly one non-zero element which is a 1 located at its $i-$th row and $j-$th column. This matrix is represented by $\Gamma(\epsilon) = \Gamma((\nu_j, \nu_i))$. Similarly, the vertex-index of any $\nu_i \in \mathcal{V}$ is defined as a $|\mathcal{V}| \times 1$ column vector with exactly one non-zero element which is a 1 located at its $i-$th row. This vector is denoted by $\sigma(\nu_i)$. The adjacency matrix of $\mathcal{G}$ is given by $\mathcal{A}(\mathcal{G}) = \sum_{\epsilon \in \mathcal{E}} \Gamma(\epsilon)$, its degree matrix is defined as $\Delta(\mathcal{G}) = \sum_{\nu \in \mathcal{V}} d^{\mathcal{G}}_{\nu} \{\nu\} \Gamma((\nu, \nu))$, and the corresponding in-degree graph Laplacian is given by $\mathcal{L}(\mathcal{G}) = \Delta(\mathcal{G}) - \mathcal{A}(\mathcal{G})$.

Given an integer $k \in \mathbb{N}$, a set of (possibly repeated)
indices $\{\alpha_1, \alpha_2, \ldots, \alpha_k\} \subseteq \mathbb{N}_{|V|}$ and two vertices $\tau, \nu \in V$, an ordered sequence of edges of the form $W := (\tau, \nu_{\alpha_1}), (\nu_{\alpha_1}, \nu_{\alpha_2}), \ldots, (\nu_{\alpha_{k-1}}, \nu_{\alpha_k}), (\nu_{\alpha_k}, \nu)$ is called a $\tau \nu$ walk with start-node $\tau$, end-node $\nu$ and length $k + 1$. A cycle on node $\nu$ signifies a $\nu \nu$ walk. If $W$ is an in-weighting on $G$, then $\omega(W, W) = \prod_{(\nu_{\alpha_i}, \nu_{\alpha_{i+1}}) \in W} |W|_{ij}$ is referred to as the weight of walk $W$ w.r.t. $W$. In the same venue, the number of $\tau \nu$ walks with length $k$, denoted by $c_k(G; \tau, \nu)$, is called the $k$-th connectivity of $\tau$ to $\nu$ in digraph $G$, and by convention, $c_0(G; \tau, \nu) = 0$ if $\tau \neq \nu$, while $c_0(G; \nu, \nu) = 1$.

Moreover, the integer
\[
d(G; \nu_j, \nu_i) = \min_{k \in K} \{k\},
\]
is referred to as the distance from $\nu_j$ to $\nu_i$ in $G$, and by convention $d(G; \nu_j, \nu_i) = \infty$ if $\forall k \in \mathbb{N}, c_k(G; \nu_j, \nu_i) = 0$. For any $\{s, p\} \subseteq \mathbb{N}_{|V|}$, $\Omega^k(G; \nu_s, \nu_p)$ is the set of all $\nu_s \nu_p$ walks in $G$ with length $k$. Similarly for $\{s, i, p\} \subseteq \mathbb{N}_{|V|}$, $\Omega^k(G; \nu_s, \nu_i, \nu_p) = \{W \in \Omega^k(G; \nu_s, \nu_p); (\nu_s, \nu_i) \in W\}$, i.e. the set of $\nu_s \nu_p$ walks with length $k$ that include the edge $(\nu_s, \nu_i)$. Given a set of walks $\Omega$ in digraph $G$ and in-weightings $W_1$ and $W_2$ on $G$, the functions:
\[
\Phi(\Omega, W_1) = \sum_{W \in \Omega} \omega(W, W_1),
\]
\[
\Psi(\Omega, W_1, W_2) = \Phi(\Omega, W_1) - \Phi(\Omega, W_2),
\]
are defined, which will find use in the proof of the main theorem in Section III that follows. It is also known for an in-weighting $W$ on $G$, and vertices $\nu_s, \nu_p \in \mathbb{V}$, that [8]:

\[
\Phi(\Omega^k(G; \nu_s, \nu_p), W) = [W^{k}]_{ps}.
\]

III. DERIVATIVES OF THE CONSENSUS RESPONSE

Given $n \in \mathbb{N}$, consider a multi-agent system comprised of a set $S = \{x_i, i \in \mathbb{N}_n\}$ of $n$ single integrator agents, where $x_i, i \in \mathbb{N}_n$, is the scalar state of agent $i$. Under the linear agreement protocol, if the interaction structure between the agents is represented by a directed information flow graph $\mathcal{G} = (\mathbb{V}, E)$, where $\mathbb{V} = \{\nu_s, i \in \mathbb{N}_n\}$ and $\forall i \in \mathbb{N}_n, \nu_i$ corresponds to $x_i$, then the dynamic evolution law for the agents is given by:

\[
x(t) = -L(\mathcal{G})x(t), \quad t > 0,
\]
where $x(t) = (x_1(t), x_2(t), \ldots, x_n(t))^T$. For an initial condition $x(0) \in \mathbb{R}^{|V|}$, the matrix exponential solution to (2) can be derived as:

\[
x(t) = e^{-L(\mathcal{G})t}x(0), \quad t \geq 0,
\]
and for a particular agent $x_p \in S$ represented by the vertex $\nu_p \in \mathbb{V}$, the temporal evolution of its state is then given by:

\[
x_p(t) = \sigma(\nu_p)^T e^{-L(\mathcal{G})t}x(0), \quad t \geq 0.
\]

The next theorem paves the way for a method to detect and isolate the failure of any links in the network and at the same instant as they fail. The latter’s significance is better understood upon noting that if the time of failure is random and has a continuous sample space, then “simultaneous” failure of more than one link is a measure zero event, hence justifying the focus of investigation in this paper, which is on the “single” link failures. It is further assumed that at each instant of time, the designer is given access to the response of a subset of agents as well as the network information flow digraph prior to the failure. In the case of detection, the designer is interested in determining the existence of any single link failure in the network at the instant of failure. For the isolation problem, however, the designer would like to determine “instantaneously”, not only the existence of a failure, but also its location. That is to determine which link, if any, has failed and exactly at the same instant as it fails.

The proposed method is based on the derivatives of the consensus response given in (3). The proof ingredients are as follows. Terms of the form $(-L(\mathcal{G}))^k$ appear upon taking the $k$-th derivative of (3). Corresponding to a digraph $\mathcal{G}$, a new digraph $\mathcal{G}$ is defined by adding a self-loop on each node and it is then noted that $-L(\mathcal{G})$ define an in-weighting on $\mathcal{G}$. Hence, the stage is set for the application of the summation formula given in (1) and the rest of the proof carries through by partitioning the set of walks over which (1) is summed.

In the theorem, $\mathcal{G}_1$ represents the original (healthy system) digraph and $\mathcal{G}_2$ is the digraph that is missing a single link. The removed link is $\epsilon := (\nu_j, \nu_i)$ and agent $x$ corresponding to vertex $\nu_p$ is an agent whose response $x(t)$ is being observed by the designer.

**Theorem 1.** Given a multi-agent system $S$ and its associated digraph $\mathcal{G}_1 = (\mathbb{V}, \mathcal{E}_1)$, consider a vertex $\nu_p \in \mathbb{V}$ corresponding to agent $x \in S$, and an edge $\epsilon := (\nu_j, \nu_i) \in \mathcal{E}_1$, and denote $\mathcal{G}_2 = (\mathbb{V}, \mathcal{E}_1 \cup \{\epsilon\})$. Starting from the same initial condition $x(0)$, for $t > 0$, let $x_{\mathcal{G}_1}(t)$ and $x_{\mathcal{G}_2}(t)$ denote the state of the agent $x$ calculated in the digraphs $\mathcal{G}_1$ and $\mathcal{G}_2$, respectively; and define $\nabla(k) := \frac{\partial}{\partial \epsilon} (x_{\mathcal{G}_2}(0^+)) - \frac{\partial}{\partial \epsilon} (x_{\mathcal{G}_2}(0^-))$. The following statements hold true: (i) $\forall k < d(\mathcal{G}_1; \nu_j, \nu_p) - 1, \nabla(k) = 0$, and (ii) for $k = d(\mathcal{G}_1; \nu_j, \nu_p)$, $\nabla(k) = c_{k-1}(\mathcal{G}_1; \nu_j, \nu_p)(x_j(0) - x_i(0))$.

**Proof.** Define $\tilde{\mathcal{G}}_1 = (\mathbb{V}, \mathcal{E}_1 \cup_{\nu \in \mathbb{V}} \{(\nu, \nu)\})$ and $\tilde{\mathcal{G}}_2 = (\mathbb{V}, \mathcal{E}_1 \cup_{\nu \in \mathbb{V}} \{(\nu, \nu)\})$ and note that $-L(\tilde{\mathcal{G}}_1)$ and $-L(\tilde{\mathcal{G}}_2)$ define proper in-weightings on $\tilde{\mathcal{G}}_1$ and $\tilde{\mathcal{G}}_2$, respectively. Differentiating both sides of (4) $k$ times yields:

\[
\nabla(k) = \sigma(\nu_p)^T ((-L(\mathcal{G}_1))^k - (-L(\mathcal{G}_2))^k)x(0),
\]
which can be rewritten as:

\[
\nabla(k) = \sum_{s=1}^{[V]} \left([-L(\mathcal{G}_1)]^k \right)_{ps} - \left([-L(\mathcal{G}_2)]^k \right)_{ps} x_s(0),
\]
where $x_s(0)$ is a scalar value that denotes the initial state of the agent corresponding to vertex $\nu_s \in \mathbb{V}$. Replacing (1) into the preceding expression of $\nabla(k)$ leads to (5) at the top of the succeeding page.

For part (i), note that if $k \leq d(\mathcal{G}_1; \nu_j, \nu_p) - 1$, there are no walks of length $k$ that include $(\nu_j, \nu_i)$ as an edge and terminate at node $\nu_p$. Hence, for $k \leq d(\mathcal{G}_1; \nu_j, \nu_p) - 1$, $\Omega^k(\tilde{\mathcal{G}}_2; \nu_s, \nu_p) = \Omega^k(\tilde{\mathcal{G}}_1; \nu_s, \nu_p)$, since $\tilde{\mathcal{G}}_1$ and $\tilde{\mathcal{G}}_2$ differ only at edge $(\nu_j, \nu_i)$ and thus they have the same set of
\[\nabla(k) = \sum_{s=1}^{|V|} \left[ \Phi(\Omega^k(\tilde{G}_1; \nu_s, \nu_p), -\mathcal{L}(G_1)) - \Phi(\Omega^k(\tilde{G}_2; \nu_s, \nu_p), -\mathcal{L}(G_2)) \right]x_s(0), \quad (5)\]

\[\nabla(k) = \sum_{s=1}^{|V|} \Psi(\Omega^k(\tilde{G}_2; \nu_s, \nu_p), -\mathcal{L}(G_1), -\mathcal{L}(G_2))x_s(0). \quad (6)\]

\[\nabla_1(k) = \sum_{s=1}^{|V|} \Psi(\Omega^k(\tilde{G}_2; \nu_s, \nu_p), \Omega^k(\tilde{G}_1; \nu_s, \nu_p), -\mathcal{L}(G_1), -\mathcal{L}(G_2))x_s(0). \quad (7)\]

\[\nabla_2(k) = \sum_{s=1}^{|V|} \Psi(\Omega^k(\tilde{G}_2; \nu_s, \nu_p), -\mathcal{L}(G_1), -\mathcal{L}(G_2))x_s(0). \quad (8)\]

\[\nabla_i(k) = \left[ \Phi(\Omega^k(\tilde{G}_1; \nu_i, \nu_p), -\mathcal{L}(G_1)) - \Phi(\Omega^k(\tilde{G}_2; \nu_i, \nu_p), -\mathcal{L}(G_2)) \right]x_i(0) \quad (9)\]

\[\nabla_j(k) = \left[ \Phi(\Omega^k(\tilde{G}_1; \nu_j, \nu_p), -\mathcal{L}(G_1)) - \Phi(\Omega^k(\tilde{G}_2; \nu_j, \nu_p), -\mathcal{L}(G_2)) \right]x_j(0) \quad (10)\]

\[\nabla(k) = \sum_{s \in N_{|V|} \setminus \{i, j\}} \left[ \Phi(\Omega^k(\tilde{G}_1; \nu_s, \nu_p), -\mathcal{L}(G_1)) - \Phi(\Omega^k(\tilde{G}_2; \nu_s, \nu_p), -\mathcal{L}(G_2)) \right]x_s(0). \quad (11)\]

walks with no \((\nu_j, \nu_i)\) edge. Given that \(\Omega^k(\tilde{G}_2; \nu_s, \nu_p) = \Omega^k(\tilde{G}_1; \nu_s, \nu_p)\), (5) can be rewritten as (6) at the top. The next step is to partition the set of walks \(\Omega^k(\tilde{G}_2; \nu_s, \nu_p)\) into two disjoint sets, \(\Omega^k(\tilde{G}_2; \nu_s, \nu_p) \setminus \Omega^k(\tilde{G}_2; \nu_s, \nu_p)\) and \(\Omega^k(\tilde{G}_2; \nu_s, \nu_p)\). Thence, \(\nabla(k)\) in (6) can be rewritten as \(\nabla(k) = \nabla_1(k) + \nabla_2(k)\), where \(\nabla_1(k)\) and \(\nabla_2(k)\) are given by (7) and (8), and they correspond to the contributions made by the walks in \(\Omega^k(\tilde{G}_2; \nu_s, \nu_p) \setminus \Omega^k(\tilde{G}_2; \nu_s, \nu_p)\) and \(\Omega^k(\tilde{G}_2; \nu_s, \nu_p)\), respectively. To finish the proof of part (i), note that \(\forall W \in \Omega^k(\tilde{G}_2; \nu_s, \nu_p) \setminus \Omega^k(\tilde{G}_2; \nu_s, \nu_p), \omega(W, -\mathcal{L}(G_1)) = \omega(W, -\mathcal{L}(G_2))\), since any such walk \(W\) includes neither of the edges \((\nu_i, \nu_j)\) and \((\nu_j, \nu_i)\), which are the only edges at which \(-\mathcal{L}(G_1)\) and \(-\mathcal{L}(G_2)\) differ. It is therefore true that \(\forall s \in N_{|V|}, \Psi(\Omega^k(\tilde{G}_2; \nu_s, \nu_p) \setminus \Omega^k(\tilde{G}_2; \nu_s, \nu_p), -\mathcal{L}(G_1), -\mathcal{L}(G_2)) = 0\), and \(\nabla_1(k) = 0\). On the other hand, from \(k \leq d(\tilde{G}_1; \nu_j, \nu_p) - 1 \leq d(\tilde{G}_1; \nu_i, \nu_p)\), which in turn implies that \(\Omega^k(\tilde{G}_2; \nu_s, \nu_p) = \emptyset\). Because with \(k \leq d(\tilde{G}_1; \nu_j, \nu_p)\) there can be no \(\nu_j \nu_p\) walks of length \(k\) in \(\tilde{G}_2\) with \((\nu_s, \nu_i)\) as an edge, for otherwise one can remove \((\nu_s, \nu_i)\) and construct a \(\nu_j \nu_p\) walk in \(\tilde{G}_1\) whose length is strictly less than \(d(\tilde{G}_1; \nu_j, \nu_p)\). Now from \(\Omega^k(\tilde{G}_2; \nu_s, \nu_i, \nu_p) = \emptyset\) it follows that \(\forall s \in N_{|V|}, \Psi(\Omega^k(\tilde{G}_2; \nu_s, \nu_i, \nu_p), -\mathcal{L}(G_1), -\mathcal{L}(G_2)) = 0\), and \(\nabla_2(k) = 0\), as well. Thus far, it is shown that \(\nabla_1(k) = \nabla_2(k) = 0\), whence \(\nabla(k) = \nabla_1(k) + \nabla_2(k) = 0\), completing the proof for \(k \leq d(\tilde{G}_1; \nu_j, \nu_p) - 1\).

For the case of \(k = d(\tilde{G}_1; \nu_j, \nu_p)\) in part (ii), first note that by conditioning on the choice of the start-node \(s\), the summation in (5) can be rewritten as: \(\nabla(k) = \nabla_i(k) + \nabla_j(k) + \nabla(k)\), where the terms are defined in (9) to (11) at the top, and they measure the contributions due to the walks starting from nodes \(\nu_j, \nu_i\), and the rest of the nodes, respectively. In the following paragraphs, the expressions for each of the above three terms are simplified in the respective order, leading to the equation in part (ii).

The first step in simplifying the expression of \(\nabla_i(k)\) is (\(\nabla_i\)), which follows by partitioning the sets \(\Omega^k(\tilde{G}_1; \nu_i, \nu_p)\) and \(\Omega^k(\tilde{G}_2; \nu_i, \nu_p)\). Next note that no \(\nu_i \nu_p\) of length \(k = d(\tilde{G}_1; \nu_j, \nu_p)\) can include \((\nu_j, \nu_i)\) as an edge. This statement is trivial in the case of \(\tilde{G}_2\). To see why it is true for \(\tilde{G}_1\), suppose the contrary that there exists a \(\nu_i \nu_p\) walk \(W_1\) of length \(k\) with \((\nu_j, \nu_i)\) as an edge. There is therefore a cycle on \(\nu_i\) whose length is at least 2, and removing that from \(W_1\) yields a new \(\nu_j \nu_p\) walk \(W_2\) with length at most \(k - 2\). Now \((\nu_j, \nu_i)W_2\) is a \(\nu_j \nu_p\) walk of length at most \(k - 1\), which is a contradiction, since


1. **Detection**
   Given a digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, propose a (preferably minimal) subset of vertices $\mathcal{M}_D \subset \mathcal{V}$ such that $\forall v \in \mathcal{E}, \exists v_p \in \mathcal{M}_D, (v_p, v) \notin \mathcal{R}_0$.

   Let $\mathcal{M}_D \subset \mathcal{V}$ be any solution to Problem 1, then a link has failed at (the arbitrarily shifted) time $t = 0$ if, and only if, there exists some $v_p \in \mathcal{M}_D$ and $k \in \mathbb{N}$ such that $\frac{d^k}{dt^k} x(p)(0^+) - \frac{d^k}{dt^k} x(p)(0^-) \neq 0$. That is, by observing the first $z$ derivatives of the responses of all the nodes in $\mathcal{M}_D$, the designer determines that a failure occurs if, and only if, a jump is observed in one or more nodes in $\mathcal{M}_D$. For the isolation problem the following extra tool comes handy. Given a subset of nodes $\mathcal{M}$ and any edge $e = (v, v')$ in $\mathcal{E}$ define the indicator set of $e$ w.r.t. $\mathcal{M}$ as the function $I : \mathcal{P}(\mathcal{V}) \times \mathcal{E} \rightarrow \mathcal{P}(\mathbb{N} \cup \{0\}) \times \mathcal{M}$, given by $I(M; e) = \{(k, v_p) \in (\mathbb{N} \cup \{0\}) \times M \mid (v_p, v) \in \mathcal{E} \}$.

   **Problem 2.** [Isolation] Given a digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, propose a (preferably minimal) subset of vertices $\mathcal{M}_I \subset \mathcal{V}$ such that $\mathcal{M}_I$ is a solution to Problem 1 and $\forall \alpha, \beta \in \mathcal{E}, \alpha \neq \beta \iff I(\mathcal{M}_I; \alpha) \neq I(\mathcal{M}_I; \beta)$.

   Given a solution $\mathcal{M}_I$ to Problem 2, any edge $e$ is uniquely determined by the set of ordered pairs $I(\mathcal{M}_I; e)$. Hence, the edge $e$ has failed at (the arbitrarily shifted) time $t = 0$ if, and only if, $\forall (k, v_p) \in I(\mathcal{M}_I; e)$, $k > 0$, $\frac{d^k}{dt^k} x(p)(0^+) - \frac{d^k}{dt^k} x(p)(0^-) \neq 0$, and $\forall (0, v_p) \in I(\mathcal{M}_I; e)$, $\forall m \in \mathbb{N}$, $\frac{d^m}{dt^m} x(p)(0^+) - \frac{d^m}{dt^m} x(p)(0^-) = 0$. In other words, by observing the first $z$ derivatives of the responses of all the nodes $v_p$ in $\mathcal{M}_I$, the designer determines that the edge $e$ has failed if, and only if, a jump is observed in the $k$-th derivative of the response of any node $v_p$ for which $(k, v_p) \in I(\mathcal{M}_I; e)$.

### IV. Placement of the Sensors

Throughout this section, $\mathcal{S} = \{x_i, i \in \mathbb{N}_n\}$ is a multi-agent system, with $n \in \mathbb{N}$ a fixed integer. The digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is associated with $\mathcal{S}$, where $\mathcal{V} = \{v_s, s \in \mathbb{N}_n\}$ and $\forall s \in \mathbb{N}_n$, $v_s$ corresponds to $x_s$. Moreover, $z \in \mathbb{N}_0$ is a fixed integer denoting the highest order of the derivatives of the responses of the consensus dynamics on $\mathcal{S}$, to which the designer has access. Similarly, with $\{k, p, i, j\} \subset \mathbb{N}_n$, the order of the derivatives is denoted by $k \in \mathbb{N}_n$, $v_p \in \mathcal{V}$ is a node whose response is being observed, $\epsilon := (v_j, v_i) \in \mathcal{E}$ is an edge whose failure at time $t = 0$ is to be detected and/or isolated by the designer, and it is further assumed that $x_i(t) = x_j(t)$, i.e. whenever an edge fails, the state values at its head and tail nodes do not coincide.

Given $k \in \mathbb{N}_n$, let $\mathcal{R}_k$ and $\mathcal{R}_0$ be binary relations between the vertices and edges of $\mathcal{G}$, such that $\forall v_p \in \mathcal{V}$ and $\forall e \in \mathcal{E}$, where $\epsilon := (v_j, v_i)$ for some $\{v_j, v_i\} \subset \mathcal{V}$, $\{v_p, \epsilon\} \in \mathcal{R}_k$ or $(d(G_1; v_j, v_i) = k) \land (d(G_1; v_j, v_i) = k - 1)$ and $(v_p, \epsilon) \in \mathcal{R}_0 \iff \forall k \in \mathbb{N}_n, (v_p, \epsilon) \notin \mathcal{R}_k$. Note that with $\epsilon = (v_j, v_i)$, by definition $(v_i, \epsilon) \in \mathcal{R}_1$ and $(v_j, \epsilon) \in \mathcal{R}_0$.

The preceding definitions are motivated by the observation made in Remark 1 that if the existence of jumps in the $k$-th derivative of the response of a node $v_p$ is to serve as an indicator for the failure of edge $(v_j, v_i)$, then it should hold true that $(d(G_1; v_j, v_i) = k) \land (d(G_1; v_j, v_i) = k - 1)$. The problems of detection and isolation are now posed and addressed as sensor placement problems in the following [9], [10].

1. **Detection**
   Given a digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, propose a (preferably minimal) subset of vertices $\mathcal{M}_D \subset \mathcal{V}$ such that $\forall e \in \mathcal{E}, \exists v_p \in \mathcal{M}_D, (v_p, e) \notin \mathcal{R}_0$.

2. **Isolation**
   Given a digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, propose a (preferably minimal) subset of vertices $\mathcal{M}_I \subset \mathcal{V}$ such that $\mathcal{M}_I$ is a solution to Problem 1 and $\forall \alpha, \beta \in \mathcal{E}, \alpha \neq \beta \iff I(\mathcal{M}_I; \alpha) \neq I(\mathcal{M}_I; \beta)$.
$k > 0$, and no jumps is observed in the first $z$ derivatives of the responses of all nodes $ν_p$ for which $(0, ν_p) ∈ T(I_M; ε)$.

In the following subsections, efficient algorithms are proposed for the selection of the observation points, which are the subsets $M_D$ and $M_I$ satisfying the requirement of the problems 1 and 2, respectively.

### A. Detection of All Failures: Coverage

For all $M ⊂ V$ define $f_D : P(V) → N_1 ∪ \{0\}$, given by $f_D(M) = |\{ε ∈ V : ν_0 ∈ M, (ν_0, ε) ∈ R_0\}|$. Note that if $\{M, M’\} ⊂ P(V)$ are such that $M ⊂ M’$ then for all $ν ∈ V$ it holds true that $f_D(M) ≤ f_D(M’)$, i.e. $f_D(\cdot)$ is a submodular set function from $P(V)$ to $N_1 ∪ \{0\}$, which satisfies the diminishing returns or discrete concavity property [10]. Routine 1 uses this property of the set function $f_D$ to compute a solution $M_D ⊂ V$ to Problem 1 using an efficient greedy heuristic [11].

#### Routine 1 Determine a Solution $M_D$ to Problem 1

**Input:** $G = (V, E)$

1. $M_D ≜ ∅$
2. while $f_D(M_D) ≠ 0$ do
3. $ν ≜ \arg\min\{f_D(M_D ∪ \{ν\}) − f_D(M_D) : \nu ∈ V \setminus M_D\}$
4. $M_D ≜ M_D ∪ \{ν\}$
5. end while

**Output:** $M_D$

#### Remark 2. The function $f_D(M_D)$ measures the coverage of set $M_D$ by counting the number of links that are not yet covered by $M_D$. At each iteration of Routine 1, the extra node $ν$ is selected and added to $M_D$ such that the number of newly covered links is maximized. Note that since for $ε = (ν_j, ν_1), (ν_i, ε) ∈ R_1$ it follows that $f_D(V) = 0$, whence Routine 1 is guaranteed to terminate.

#### B. Isolation of All Failures: Resolution

Similarly to the previous section, for all $M ⊂ V$ define $f_I : P(V) → N_1 ∪ \{0\}$, given by $f_I(M) = |\{ε ∈ V : ε ∈ E \setminus V(M; ε) = I(\{M; i\})\}|$, and note that $-f_I(\cdot)$ is also submodular. Thus, a greedy heuristic similar to the one in Subsection IV-A can be used to compute a solution $M_I ⊂ V$ to Problem 2, as follows.

#### Remark 3. The function $f_I(M_I)$ measures the resolution of set $M_I$ by counting the number of links that are not uniquely identified through their relations $R_0$ with the vertices of set $M_I$. At each iteration of Routine 2, the extra node $ν$ is selected and added to $M_I$ such that the resultant improvement in the resolution of $M_I$ is maximized. Note that unlike Problem 1, it is possible for Problem 2 to have no solutions at all, in which case Routine 2 returns $∅$. This occurs if, and only if, $f_I(V) ≠ 0$.

### V. Examples and Discussions

In the sequel, $V = \{ν_i, i ∈ N_5\}$ is a set of five vertices, where for all $i ∈ N_5$, vertex $ν_i$ corresponds to agent $x_i$ in a multi-agent system $S = \{x_i, i ∈ N_5\}$. It is further assumed that if a link $(ν_j, ν_1), \{i, j\} ∈ N_5$ fails at $t = t_f$, then $x_i(t_f) ≠ x_j(t_f)$.

![Fig. 1. A start digraph of size five and a directed cycle of length five are depicted in (a) and (b), respectively.](image)

As the first example, consider the case of a start network (Fig. 1(a)), where there are four edges in the network and all of them share the same head vertex $ν_5$. In this case, the designer can detect the failure of any single edge in the network by observing the first derivative of the response of $ν_5$; however, there are no subset of nodes that can be observed for isolating the failed edge. In fact $f_I(V) = 4$ in the case of a star network, and every edge of the network is in the same relations $R_1$ with the node $ν_5$ and $R_0$ with the rest of the nodes.

As the second example, consider a cycle (Fig. 1(b)), whose edge set is given by $E = \{ε_i, i ∈ N_5\}$, where for $i ∈ N_5 \setminus \{1\}$, $ε_i = (ν_{i-1}, ν_i)$ and $ε_5 = (ν_1, ν_5)$. Associate with every vertex $ν_i$ a row vector $d_i$ with five columns, whose $j$–th element $[d_i]_{j, j} ∈ N$ is equal to $k$ if $(ν_i, ε_j) ∈ R_k$ for some $k ∈ N_5 \cup \{0\}$. Let $D ∈ R^{5 × 5}$ be the matrix whose $i$–th row is equal to $d_i$ for all $i ∈ N_5$. Then $D$ is given by:

$$
D = \begin{pmatrix}
1 & 0 & 4 & 3 & 2 \\
2 & 1 & 0 & 4 & 3 \\
3 & 2 & 1 & 0 & 4 \\
4 & 3 & 2 & 1 & 0 \\
0 & 4 & 3 & 2 & 1
\end{pmatrix}
$$
It is evident from matrix $D$ that any two distinct vertices will offer a solution $MI$ to Problem 1, since the locations of 0 entries do not overlap for distinct vertices. Thence, the designer can detect the failure of any links in the cycle network by observing the jumps in the first four derivatives of any two nodes in the network. In the same vein, any set of two distinct vertices can also be used to uniquely determine which link has failed based on the observed jumps in the first four derivatives. For instance, taking $M_1 = \{v_2, v_3\}$, $e_2$ is the only edge whose failure will produce a jump in the first and second derivatives of $x_2(t)$ and $x_3(t)$, respectively, at the time of failure, $t = t_f$. Figs 2 and 3 depict the responses of the second and third agents as well as their derivatives, for a directed cycle of length five initialized at $x(0) = (1, 2, 3, 4, 5)^T$, where for all $t \in \mathbb{R}$, $x(t) = (x_1(t), x_2(t), x_3(t), x_4(t), x_5(t))^T$. The edge $e_2$ is removed at time $t_f = 5$, and as a result, the first derivative of $x_2(t)$ and the second derivative of $x_3(t)$ exhibit jump discontinuities at $t_f = 5$.

The above generalizes for finite cycles and stars of arbitrary size. In particular, there is no solution to the isolation problem for any star network and detection can be achieved by observing the first derivative of the common head vertex. For a cycle on the other hand, if all derivatives up to one less than the the network size are observed, then any two nodes offer a solution to not only the detection, but also the isolation problem.

VI. CONCLUSIONS

In this paper, a method was developed, both analytically and algorithmically, that enables the designer of a multi-agent system to detect and isolate single link failures, based on the observed jumps in the derivatives of the output responses of a subset of nodes under the Laplacian consensus dynamics. A theorem was presented that relates the jumps in the derivatives at the time of failure to the distance of the the failed links from the observation points, which if appropriately selected would then enable the designer to infer the failed link based on the observed jumps. In the proof of the theorem, the graph Laplacian can be replaced with any well-defined in-weighting on the graph, so that an extension to the general (non-Laplacian) linear dynamics is possible. In the future, the authors hope to formalize this extension for a network of linear single-integrator agents.

REFERENCES