Distributed Fault Detection and Isolation Resilient to Network Model Uncertainties

André Teixeira, Iman Shames, Henrik Sandberg, Karl H. Johansson

***Abstract*—The ability to maintain state awareness in the face of unexpected and unmodeled errors and threats is a defining feature of a resilient control system. Therefore, in this paper we study the problem of distributed fault detection and isolation (FDI) in large networked systems with uncertain system models. The linear networked system is composed of interconnected subsystems and may be represented as a graph. The subsystems are represented by nodes, while the edges correspond to the interconnections between subsystems. Considering faults that may occur on the interconnections and subsystems, as our first contribution we propose a distributed scheme to jointly detect and isolate faults occurring in nodes and edges of the system. As our second contribution, we analyze the behavior of the proposed scheme under model uncertainties caused by the addition or removal of edges. Additionally, we propose a novel distributed FDI scheme based on local models and measurements that is resilient to changes outside of the local subsystem and achieves both fault detection and isolation. Our third contribution addresses the complexity reduction of the distributed FDI method by characterizing the minimum amount of model information and measurements needed to achieve FDI and by reducing the number of monitoring nodes. The proposed methods can be fused to design a scalable and resilient distributed FDI architecture that achieves local FDI despite unknown changes outside the local subsystem. The proposed approach is illustrated by numerical experiments on the IEEE 118-bus power network benchmark.**

***Index Terms*—Fault diagnosis, networked control systems, multi-agent systems, power systems.**

I. Introduction

Critical infrastructures such as power grids, water distribution networks, and transport systems are examples of networked systems that consist of large-scale physical processes monitored and controlled over a heterogeneous set of communication networks and computers. Although the use of such powerful software systems typically adds efficiency, flexibility, and scalability, it also increases the vulnerability to mistakes from human operators, failures in equipment, and cyber attacks against the IT infrastructure [1]–[3]. Several major incidents have been reported in the past few years. For example, the extent of the US Eastern blackout in 2003 has been blamed on malfunctioning monitoring systems [4]. Other examples include cyber security breaches recently announced [5], [6]. For these reasons the area of *resilient control systems* has emerged [3]. A major feature of a resilient control system is an ability to maintain state awareness and acceptable performance under unexpected faults and malicious attacks. It is in the light of these developments that this paper introduces new methods to localize faulty and misbehaving components in large-scale control systems.

A holistic approach to security and resilience of networked control systems is important because of the complex coupling between the physical process and the distributed software system. Unfortunately a theory for such system security is lacking. Increasing the cyber security by adding encryption and authentication schemes helps to prevent some attacks by making them harder to succeed but it would be a mistake to rely solely on such methods, as it is well-known that the overall system is not secured because some of its components are. One way to enhance resiliency of networked control systems is to design control algorithms that are robust to the effects of certain categories of faults and attacks [7]–[10]. Another way is to develop monitoring schemes to detect anomalies in the system caused by attacks and faults [11]. The latter approach in general allows faster and more effective responses to anomalies as opposed to the former, since properties of the fault such as location and fault signal can be obtained. Moreover, monitoring schemes can also improve the state-awareness of the system [12].

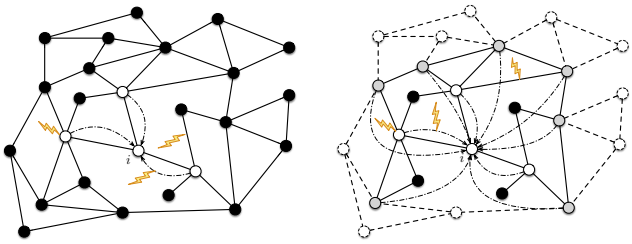
This paper focuses on the design of resilient systems using fault detection and isolation (FDI) for distributed monitoring of a network of interconnected systems. In large-scale networked systems, even benign disturbances such as model changes or unmeasured signals may hinder the detection of faults. Additionally, a global model of the system may not be available, or the large size of the system may lead to computationally intractable monitoring schemes. Hence in order to meet the demands of resilient control system components, monitoring schemes need to be architectured and designed to provide scalable solutions suitable for large-scale highly uncertain networked systems. Therefore our proposed distributed FDI scheme is resilient to model changes and external faults, not requiring the exact global model of the network to be known to the nodes.

A. Related work

There are various ways to detect and isolate a fault in a dynamical system [13]–[16]. A recent survey of different techniques can be found in [17]. One approach is to use the system model to design a set of parity equations. In the case

This work was supported in part by the Swedish Research Council under Grant 2009-4565 and Grant 2013-5523, in part by the Swedish Foundation for Strategic Research, in part by the Knut and Alice Wallenberg Foundation, in part by the University of Melbourne under the Early Career Researcher Grant, and in part by a McKenzie Fellowship.

A. Teixeira, H. Sandberg, K. H. Johansson are with the ACCESS Linnaeus Centre, KTH Royal Institute of Technology, Stockholm, Sweden. I. Shames is with the Department of Electrical and Electronic Engineering, University of Melbourne, Australia. {andretei,hsan,kallej}@kth.se, iman.shames@unimelb.edu.au



(a) (b)

Fig. 1. The networked system with faults, where nodes correspond to dynamical subsystems and undirected edges represent coupled dynamics between nodes. In distributed FDI schemes, node i aims at detecting and isolating faults on the solid white nodes and edges incident to them. Scenario (a) depicts the case where node i has access to measurements from its neighbors, represented by directed edges, and knows the entire network model. In scenario (b), node i only knows a local model of the network, where the dashed nodes and edges are unknown to node i. Moreover, node i receives measurements from the solid white and gray nodes.

of dynamical systems, such parity equations can be obtained by exploiting the temporal correlation among state, input, and output variables for a given time-horizon. This approach was used in [18] to design a centralized FDI scheme insensitive to certain model changes and disturbances. Our approach is similar, but relies on an observer-based approach and results in a distributed FDI scheme.

Observer-based FDI approaches have been well studied and some of these methods have been proposed for power systems [19], [20]. However, distributed FDI for systems comprised of a network of autonomous nodes is still in its infancy. Recently a distributed FDI scheme for a network of interconnected first-order systems was proposed [21]. The authors analyzed limitations on fault detectability and isolability in a system theoretic perspective. A similar distributed FDI scheme for interconnected second-order systems was proposed in [22]. In both contributions, the exact model of the system is assumed to be known. Distributed FDI schemes using uncertain models were proposed in [23]. However, these schemes require bounded interconnections between the subsystems and knowledge of these bounds. A similar approach was followed by [24] and applied to nonlinear power system models, but in addition to bounded model uncertainty they required also communication between neighboring FDI filters.

B. Contributions

This paper tackles the problem of distributed FDI for large-scale interconnected systems with respect to different fault models. The networked system with different fault types are illustrated in Fig. 1. The networked system is composed of interconnected individual subsystems, represented by nodes. Each node has access to local measurements from nodes in its vicinity, represented by directed edges. As an example, the measurements available to node i are depicted in Fig. 1. The interconnections between subsystems are represented by undirected edges between nodes and model either physical couplings, as in the case of power networks, or distributed control laws computed based on the local measurements, which are present, for instance, in mobile multi-agent systems. Faults may affect the network through the nodes, undirected edges, and directed edges. Given the system model and local measurements, distributed FDI aims at having each node of the network detecting and isolating faults in its vicinity, as illustrated in Fig. 1.

First we tackle the problem of distributed FDI with respect to faulty nodes and faulty edges. The proposed schemes extend the work in [22], which addressed the distributed FDI problem for faulty nodes. In particular, we consider schemes based on Unknown Input Observers (UIO) and, given the local measurements and system model as depicted in Fig. 1(a), we derive results on the existence of UIOs at each node for the different fault models.

As our second contribution, we consider the case where the UIOs are designed based on uncertain network models. More precisely, the model uncertainty is caused by the removal of edges or nodes with respect to the nominal model. The proposed distributed FDI scheme is shown to be somewhat resilient to network changes that are external to a node's local subsystem, i.e. that occur on the dashed nodes or edges in Fig. 1(b). Additionally, we propose a novel distributed FDI scheme based on local models and an augmented set of measurements from the local subsystem, as illustrated in Fig. 1(b). As opposed to approaches similar to [23], [24], bounding the subsystems' interactions is not required. Instead, by using the additional measurements, the local FDI filter can be decoupled from faults and model changes in the external subsystems and it can detect and isolate faults in the neighboring nodes.

Our third contribution is to address the complexity reduction of the distributed FDI scheme. More precisely, leveraging on our second contribution, we outline the minimum amount of model information and measurements that are sufficient for a node to achieve FDI using only its local measurements and models. In particular, our results show that using the local model from a node's 2-hop neighborhood and the corresponding measurements may not be optimal. The proposed scheme has reduced computational complexity and required model knowledge compared to the schemes such as [10], [21], [22], which use the global system's model. Moreover, we propose a method to reduce the number of monitoring nodes while ensuring that all nodes are being monitored. Importantly, we do not assume that the monitoring nodes exchange information with each other.

C. Outline

The outline of the paper is as follows. In Section II we describe the system and fault models and define the problem of distributed FDI. The distributed FDI scheme for faulty nodes and edges is detailed in Section III. In Section IV we show how to distributedly detect faults when the network model is uncertain using two different methods. The first method adapts the detection thresholds of the original distributed FDI, while the second consists of a novel distributed FDI method based on local models that not only requires less computation

than the one presented in Section III, but also is capable of handling uncertain network models. In Section V we propose methods to reduce the computational burden of the methods described in Sections III. Some numerical examples are given in Section VI. Concluding remarks are presented in the last section.

II. Networked Control System

Consider a network of N interconnected dynamical systems and let $\mathcal{G}(\mathcal{V},\mathcal{E})$ be the underlying graph of this network, where $\mathcal{V}=\{i\}_{i=1}^{N}$ is the vertex set and $\mathcal{E}\subseteq\mathcal{V}\times\mathcal{V}$ is the edge set of the graph. Denote $\mathcal{A}\in\mathbb{R}^{N\times N}$ as the weighted adjacency matrix with nonnegative entries. The undirected edge $\{i,j\}$ is incident to vertices i and j if nodes i and j share a communication link, in which case the corresponding entry in the adjacency matrix $[\mathcal{A}]_{ij}$ is positive. The degree of node i is $\deg(i)=\mathcal{A}\mathbf{1}_N=\sum_{j\in\mathcal{N}_i}[\mathcal{A}]_{ij}$, where the entries of $\mathbf{1}_N\in\mathbb{R}^N$ are equal to 1, $\mathcal{N}_i=\{j\in\mathcal{V}:\{i,j\}\in\mathcal{E}\}$ is the neighborhood set of i with $N_i\triangleq|\mathcal{N}_i|$, and the degree matrix of $\mathcal{G}$ is $\Delta\triangleq\text{diag}(\deg(1),\dots,\deg(N))$. The Laplacian of $\mathcal{G}$ is defined as $\mathcal{L}(\mathcal{G})=\Delta-\mathcal{A}$. Consider a subset of the vertex set $\tilde{\mathcal{V}}\subseteq\mathcal{V}$ and a subset of the edge set $\tilde{\mathcal{E}}\subseteq\mathcal{E}$. The subgraph of $\mathcal{G}$ induced by $\tilde{\mathcal{V}}$ and $\tilde{\mathcal{E}}$ is denoted as $\tilde{\mathcal{G}}(\tilde{\mathcal{V}},\tilde{\mathcal{E}})$. Moreover, assume that the state of each node is given by $x_i(t)\in\mathbb{R}^2$.

We call the set $\mathcal{N}_i^\ell\subset\mathcal{V}$ *the ℓ-hop neighbor set of node i* where $v\in\mathcal{N}_i^\ell$ if there is a path of length at most ℓ between i and v. Defining $\mathcal{V}_i^\ell=\{i\}\cup\mathcal{N}_i^\ell$, we call the subgraph $\mathcal{G}_i^\ell\left(\mathcal{V}_i^\ell,\mathcal{E}_i^\ell\right)\subseteq\mathcal{G}(\mathcal{V},\mathcal{E})$ *the ℓ-hop neighborhood graph of node i* where $\{v,u\}\in\mathcal{E}_i^\ell$ if $\{v,u\}\in\mathcal{E}$ and $u,v\in\mathcal{N}_i^\ell$. For the case where $\ell=1$, we drop the superscript for the ease of notation. We call the graph $\mathcal{P}_i(\mathcal{V}_{P_i},\mathcal{E}_{P_i})\subseteq\mathcal{G}(\mathcal{V},\mathcal{E})$, where $\mathcal{V}_{P_i}=\{i\}\cup\mathcal{N}_i\cup\overline{\mathcal{N}}_i$, and $\mathcal{E}_{P_i}=\mathcal{E}_i\cup\overline{\mathcal{E}}_i$, *the proximity graph of node i* where $\{v,u\}\in\mathcal{E}_i$ if $\{v,u\}\in\mathcal{E}$ and $u,v\in\mathcal{N}_i$. Moreover, $\overline{\mathcal{N}}_i$ is the set of all the nodes in the network that are not in $\mathcal{N}_i$ but share a link with at least one of the nodes in $\mathcal{N}_i$, and $\overline{\mathcal{E}}_i$ is the set of all edges incident to at least one of the nodes in $\mathcal{N}_i$ that are not in $\mathcal{E}_i$. Examples for the notation above are given in Fig. 2.

In this paper we consider linear time-invariant networked systems described by

$$\begin{aligned}\dot{x}(t)&=Ax(t)+Bv(t)+Ef(t),\\ y_i(t)&=C_ix(t)+D_if(t),\quad\forall\, i\in\mathcal{V},\end{aligned}\tag{1}$$

where $x(t)\in\mathbb{R}^n$ is the global state vector containing all the agents' states, $v(t)\in\mathbb{R}^N$ is a known input vector, $y_i(t)\in\mathbb{R}^{m_i}$ is the set of measurements available at node i, and $f(t)\in\mathbb{R}^p$ is an unknown vector of faults affecting the system. We are interested in the problem of distributed fault detection and isolation, as described below.

Definition 1 (Distributed fault detection and isolation)**.** *Consider the system* (1) *and suppose each node i has a model of the system and a local set of measurements $y_i(t)$ to design a FDI scheme. A fault $f(t)\not\equiv 0$ is said to be* detected *if at least one node $i\in\mathcal{V}$ decides that there exists an active fault in the network. Furthermore, a fault is said to be* isolated *if there exists a set of nodes that detect the fault and identify the faulty components, i.e. identify the non-zero elements of $f(t)$.*

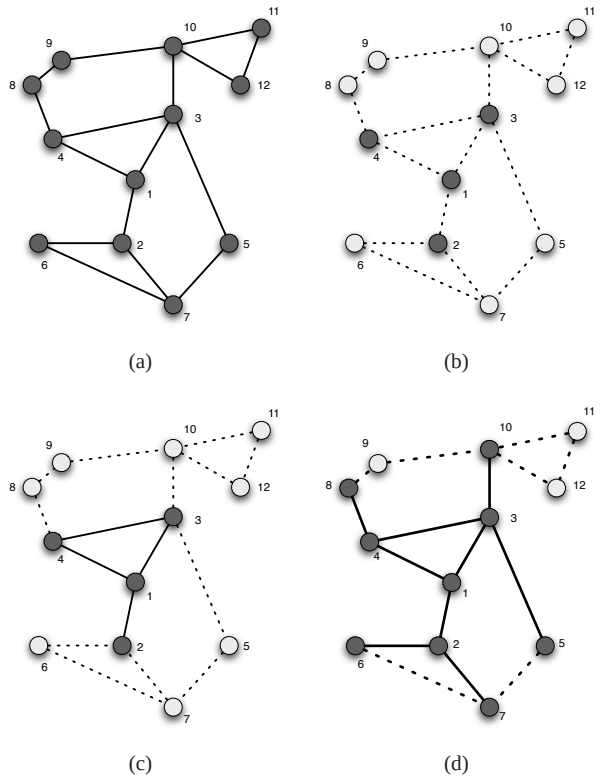


Fig. 2. (a) A network with 12 nodes. (b) The set of one-hop neighbors of node 1, $\mathcal{N}_1$, are nodes $\{2,3,4\}$ and are coloured darker. (c) The one-hop neighborhood graph of node 1, $\mathcal{G}_1$, is the set of dark nodes connected by solid lines. (d) The graph represented by dark nodes that are connected to each other by solid lines is the proximity graph of node 1, i.e., $\mathcal{P}_1$.

The main aim of this work is to leverage the structural properties of the networked system (1) to characterize under what conditions the problem of distributed fault detection and isolation can be solved. In particular, we focus on the networked second-order systems, while similar results for networked first-order systems can be obtained, see for instance [21], [25]. For this case the state of each node, $x_i(t)=[\xi_i(t)\ \zeta_i(t)]^\top$, $\xi_i(t)$, and $\zeta_i(t)\in\mathbb{R}$, is governed by

$$\dot{\xi}_i(t)=\zeta_i(t)\tag{2a}$$
$$\dot{\zeta}_i(t)=u_i(t)+v_i(t)+f_i(t),\tag{2b}$$

where $\xi_i(t)$ and $\zeta_i(t)$ are the scalar states, $v_i(t)$ is the i-th entry of the external reference input $v(t)$, $u_i(t)$ is a scalar distributed control input capturing the interactions between neighboring nodes, and $f_i(t)$ is an unknown fault affecting node i. Additionally, each agent i has access to its own states and receives measurements of its neighbors' states, possibly corrupted. Denoting $x(t)=[\xi_1(t)\ \dots\ \xi_N(t)\ \zeta_1(t)\ \dots\ \zeta_N(t)]^\top$ as the global system state, the measurement vector with corrupted measurements is described as

$$y_i(t)=C_ix(t)+C_i\sum_{j\in\mathcal{N}_i}\left(l_jf_{ij}^{\xi}(t)+l_{N+j}f_{ij}^{\zeta}(t)\right),\tag{3}$$

where $j_k\in\mathcal{N}_i$ for all $k=1,\dots,N_i$, $l_i\in\mathbb{R}^{2N}$ is the i-th

column of I_{2N}, and $C_i = [\bar{C}_i^\top \ \bar{C}_i^\top]^\top$, with $\bar{C}_i \in \mathbb{R}^{|\mathcal{V}_i^1| \times N}$ being a full row rank matrix where each of the rows have all zero entries except for one entry at the j-th position that corresponds to those nodes that are in $\mathcal{V}_i^1 = \{i\} \cup \mathcal{N}_i$. The variables $f_{ij}^\xi(t)$ and $f_{ij}^\zeta(t)$ for $j \in \mathcal{N}_i$ denote measurement corruptions on ξ_j and ζ_j, respectively.

The distributed control input $u_i(t)$ is given by the linear control law on $y_i(t)$:

$$u_i(t) = \sum_{j \in N_i} (w_{ij} + f_{ij}^w(t)) \Big[(\xi_j(t) + f_{ij}^\xi(t) - \xi_i(t)) + \mu(\zeta_j(t) + f_{ij}^\zeta(t) - \zeta_i(t)) \Big] - \kappa_i \zeta_i(t), \tag{4}$$

where $w_{ij} = w_{ji} \in \mathbb{R}_{>0}$ are the edge weights, $\kappa_i, \mu \in \mathbb{R}_{\geq 0}$ for $i, j = 1, \ldots, N$, and $f_{ij}^w(t) = f_{ji}^w(t)$ is an unknown fault affecting the weight of the edge $\{i, j\}$.

The overall dynamics of the networked system under the control law (4) are described by (1) with

$$A = \begin{bmatrix} 0_N & I_N \\ -\mathcal{L} & -\mu\mathcal{L} - \overline{K} \end{bmatrix}, \quad B = \begin{bmatrix} 0_N \\ I_N \end{bmatrix}. \tag{5}$$

The matrix $\mathcal{L}$ is the weighted Laplacian matrix associated with the network where w_{ij} is the weight of edge $\{i, j\}$, and $\overline{K} = \text{diag}(\kappa_1, \ldots, \kappa_N)$.

Given the global system model (1), the node dynamics (2), the local measurements (3), and the distributed control law (4), we define faulty nodes and faulty edges as follows.

Definition 2. *A node $i \in \mathcal{V}$ is faulty if $f_i(t) \not\equiv 0$. The system affected by the fault $f(t) = f_i(t)$ is modeled by* (1) *with $E = b_i$ and $D_i = 0$, where b_i is the i-th column of B.*

Definition 3. *An edge $\{i, j\} \in \mathcal{E}$ is faulty if any of the signals $f_{ij}^w(t)$, $f_{ji}^w(t)$, $f_{ij}^\xi(t)$, $f_{ji}^\xi(t)$, $f_{ij}^\zeta(t)$, and $f_{ji}^\zeta(t)$ are not identically zero. Moreover, we classify edge faults as either sensing faults or parameter faults.*

1) *A fault on edge $\{i, j\}$ is a sensing fault from j to i if any of the signals $f_{ij}^\xi(t)$ and $f_{ij}^\zeta(t)$ are not identically zero and $f_{ij}^w(t) \equiv 0$. The system affected by the fault $f(t) = [f_{ij}^\xi(t) \ f_{ij}^\zeta(t)]^\top$ is modeled by* (1) *with $E = b_i[w_{ij} \ \mu w_{ij}]$ and $D_i = C_i[l_j \ b_j]$, where l_j is the j-th column of I_{2N}.*
2) *A fault on edge $\{i, j\}$ is a parameter fault if the signals $f_{ij}^\xi(t)$, $f_{ij}^\zeta(t)$, $f_{ji}^\xi(t)$, and $f_{ji}^\zeta(t)$ are identically zero and $f_{ij}^w(t) = f_{ji}^w(t) \not\equiv 0$. The system affected by the fault $f(t) = \delta_{ij}(t) f_{ij}^w(t)$ with $\delta_{ij}(t) = \xi_j(t) - \xi_i(t) + \mu(\zeta_j(t) - \zeta_i(t))$ is modeled by* (1) *with $E = b_i - b_j$ and $D_i = 0$.*

The control law described by (4) with $f(t) \equiv 0$ is a generalized form of the two following well-known control laws:

$$u_i^1(t) = -\kappa_i \zeta_i(t) + \sum_{j \in \mathcal{N}_i} w_{ij}(\xi_j(t) - \xi_i(t)), \tag{6}$$

$$u_i^2(t) = \sum_{j \in \mathcal{N}_i} w_{ij} \left[(\xi_j(t) - \xi_i(t)) + \mu(\zeta_j(t) - \zeta_i(t)) \right] \tag{7}$$

Analysis of these control laws and design rules for κ_i, w_{ij}, and μ may be found in [26], [27].

Remark 1. *Under both these control laws with $f(t) \equiv 0$, for all $i, j \in \mathcal{V}$ we have $|\xi_i - \xi_j| \to 0$ and $|\zeta_i - \zeta_j| \to 0$ exponentially fast [26], [27]. Furthermore, we denote the consensus equilibria as $\bar{x} = [\bar{\xi} \ \bar{\zeta}]^\top \otimes \mathbf{1}_N$ with $\bar{\xi} = \lim_{t \to +\infty} \xi_i(t)$ and $\bar{\zeta} = \lim_{t \to +\infty} \zeta_i(t)$, where $\otimes$ denotes the Kronecker product.*

The introduced networked system can represent many practical systems, which may lead to different edge fault models. In this paper we consider two application examples, namely mobile multi-agent systems and electric power networks. For a mobile multi-agent system [26], each node i represents a vehicle where the variables ξ_i and ζ_i can be interpreted as the corresponding position and velocity, respectively, while the edges map to communication or sensing links between the vehicles. For this system, each node implements the control law by obtaining state measurements from the neighbors, where faults in the measurements appear as sensing faults on edges, as discussed in Definition 3.1.

In the context of synchronous power systems [28], each node i is a generator or motor with ξ_i and ζ_i being the corresponding phase and frequency, respectively, and the edges represent physical transmission lines between electrical devices. In this case the control law corresponds to the model of the physical coupling between the nodes, thus being part of the physical system itself. Moreover, faults on the edges represent are actually faults on the transmission lines. In this paper, we consider that such faults correspond to changes in the transmission line parameters, namely the edge weights $w_{ij} = w_{ji}$ are affected by a fault and become $w_{ij} + f_{ij}^w(t) = w_{ji} + f_{ji}^w(t)$, corresponding to parameter faults as per Definition 3.2.

III. Distributed Fault Detection and Isolation

In this section we address the problem of distributed fault detection and isolation of faulty nodes and faulty edges. First we revisit some of the results on distributed FDI for faulty nodes derived in [22], which is later extended to the case of faulty edges.

A. Distributed FDI for Faulty Nodes

Recall the problem of distributed FDI as per Definition 1, where each node i monitors its neighborhood to detect and isolate faulty components. In the present subsection, we address the previous problem in the case of faulty nodes.

Given the control input (4) and local measurements from its neighbors (3), node i cannot compute each neighbor's input. Therefore, FDI based solely on individual models (2) is infeasible, as the neighbors trajectories cannot be estimated. However, the control inputs and corresponding trajectories can be estimated by using the global model of the networked system (1), as described next.

For each node $i = 1, \ldots, N$, consider a model of the form:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bv(t) + \sum_{k \in \mathcal{N}_i} E_k f_k(t), \\ y_i(t) &= C_i x(t) + \sum_{k \in \mathcal{N}_i} D_{i,k} f_k(t), \end{aligned} \tag{8}$$

where, recalling Definition 2, a faulty node k is modeled by $E_k = b_k$ and $D_{i,k} = 0$. For the ease of notation, in this paper we assume that there is at most one faulty node[1].

To achieve distributed FDI, we consider a scheme where each node $i \in \mathcal{V}$ constructs a bank of N_i observers. In particular, for each $k \in \mathcal{N}_i$, an observer decoupled from E_k and $D_{i,k}$ is implemented, as described next. Given the model (8), let $\hat{x}_k^i(t)$ denote the state estimate decoupled from a faulty node k and calculated by node i using the state observer

$$\begin{aligned} \dot{z}_k^i(t) &= F_k^i z_k^i(t) + T_k^i B v(t) + K_k^i y_i(t) \\ \hat{x}_k^i(t) &= z_k^i(t) + H_k^i y_i(t), \end{aligned} \tag{9}$$

where $z_k^i(t) \in \mathbb{R}^{2N}$ is the observer's state. An unknown input observer (UIO) decoupled from a faulty node k is defined as follows [16].

Definition 4. *Consider the dynamical system* (8) *and the observer* (9)*. The observer is a* UIO decoupled from a faulty node k *if* $\lim_{t\to+\infty} \|x(t) - \hat{x}_k^i(t)\| = 0$ *for any fault* $f_k(t)$.

For the observer (9) to be a UIO, the observer matrices should be designed to achieve decoupling from the faulty node k and should ensure the stability of the observer. By choosing the matrices $F_k^i, T_k^i, K_k^i, H_k^i$ to satisfy the conditions

$$\begin{aligned} &F_k^i = (A - H_k^i C_i A - K_k'^i C), \quad T_k^i = (I - H_k^i C_i) \\ &K_k^i = K_k'^i + K_k''^i, \quad K_k''^i = F_k^i H_k^i, \quad (H_k^i C_i - I)E_k = 0, \end{aligned} \tag{10}$$

where F_k^i is Hurwitz and recalling the model (8), we have the estimation error dynamics

$$\dot{e}_k^i(t) = F_k^i e_k^i(t) - T_k^i \sum_{m\in\mathcal{N}_i\backslash\{k\}} E_m f_m(t) \tag{11}$$

with $e_k^i(t) = x(t) - \hat{x}_k^i(t)$. Clearly, the error dynamics (11) do not depend on $f_k(t)$ and are stable, thus complying with Definition 4. In general, the UIO existence condition are as follows [14].

Proposition 1. *For the system* (8)*, there exists a UIO decoupled from a faulty node k in the sense of Definition 4 if and only if the following conditions hold*

$$\begin{aligned} \mathit{rank}(C_i E_k) &= \mathit{rank}(E_k) \\ \mathit{rank}\begin{bmatrix} sI - A & E_k \\ C_i & 0 \end{bmatrix} &= n + \mathit{rank}(E_k), \end{aligned} \tag{12}$$

for all $s \in \mathcal{C}$ with non-negative real parts.

Remark 2. *The UIO existence conditions* (12) *correspond to the necessary and sufficient conditions for asymptotic estimation of the unknown input $f_k(t)$. Consider the fault signal estimate $\hat{f}_k^i(t) = V(\dot{y}_i(t) - CA\hat{x}_k^i(t))$ with $V = (C_i E_k)^\dagger$ as the pseudo-inverse of $C_i E_k$. From [16, Theorem 14.4], when $y(t)$ and $\dot{y}(t)$ are available, the necessary and sufficient conditions for $\lim_{t\to+\infty} |f_k(t) - \hat{f}_k^i(t)| = 0$ are the same as the UIO existence conditions in Proposition 1.*

[1]This assumption is not essential and can be relaxed. In particular, one may take any combination of simultaneous faults and consider it as a higher-dimensional fault signal. For instance, a simultaneous fault on nodes j and k could be modeled using (8) by replacing $E_k f_k(t)$ with $[E_k\ E_j][f_k(t)\ f_j(t)]^\top$.

The UIO error dynamics (11) are driven by the j-th fault, for some $j \neq k$, if $T_k^i E_j \neq 0$. In fact, having $T_k^i E_j \neq 0$ for all $j \in \mathcal{N}_i\backslash\{k\}$, for all $k \in \mathcal{N}_i$, plays an important role in the detection and isolation logic later described. This condition can be incorporated in the UIO design, as stated by the following results.

Proposition 2. *Given the system* (8)*, suppose the UIO existence conditions* (12) *hold for a given $k \in \mathcal{N}_i$. There exists a UIO decoupled from a faulty node k with $T_k^i E_j \neq 0$ for all $j \in \mathcal{N}_i\backslash\{k\}$ if $\mathit{rank}(C_i[E_k\ E_j]) = \mathit{rank}([E_k\ E_j]) > \mathit{rank}(E_k)$, for all $j \in \mathcal{N}_i\backslash\{k\}$.*

Proof. The desired UIO must satisfy (10) and $T_k^i E_j \neq 0$ for all $j \in \mathcal{N}_i\backslash\{k\}$. Recalling that $T_k^i = (I - H_k^i C_i)$, we then have that $T_k^i E_k = 0$ and $T_k^i E_j \neq 0$ must hold. The rank condition in the proposition's statement ensures that $H_k^i = E_k \left((C_i E_k)^\top C_i E_k\right)^{-1} (C_i E_k)^\top$ satisfies $T_k^i E_k = 0$ and $T_k^i E_j \neq 0$ for all $j \in \mathcal{N}_i\backslash\{k\}$, since E_k and E_j are orthogonal. The rest of the proof follows directly from the UIO design method detailed in [14], which constructs a UIO satisfying (10) with H_k^i as chosen above. □

Given the conditions in Proposition 1, we observe that the rank condition in Proposition 2 holds when there exist UIOs for all $k \in \mathcal{N}_i$ and every pair of fault directions E_k and E_j with $j \neq k$ is linearly independent. Since the latter holds for both node and edge faults, in the remainder of the paper we focus only on the UIO existence conditions from Proposition 1. In particular, we derive results of existence and nonexistence of UIOs for the interconnected system (1) under different fault models by using the conditions of Proposition 1.

For the moment, suppose that there exists a bank of UIOs at node i, where each UIO is decoupled from a faulty node $k \in \mathcal{N}_i$. The bank of UIOs computes a set of state estimates $\hat{x}_j^i(t)$ for $j \in \mathcal{N}_i$ given the model of the system (8), which is assumed to be accurate. Intuitively, recalling that noise is neglected, a mismatch between the estimated and actual state trajectory of the system would indicate the presence of faults in the system. In fact, node i can detect faults by analyzing the difference between the estimated outputs $\hat{y}_j^i(t) = C_i \hat{x}_j^i(t)$ for all $j \in \mathcal{N}_i$ and the actual measurements $y_i(t)$, which are denoted as residual signals.

Definition 5. *The signal $r_j^i(t) \triangleq y_i(t) - C_i \hat{x}_j^i(t) = C_i e_j^i(t)$ is a residual if $\|r_j^i(t)\| = 0$ is equivalent to $\|f_k(t)\| = 0$ for all $k \neq j \in \mathcal{N}_i$.*

Note that the residual dynamics of $r_k^i(t)$ are driven by the j-th fault if $T_k^i E_j \neq 0$, which can be ensured for $j \in \mathcal{N}_i\backslash\{k\}$ through Proposition 2. Therefore, according to Definition 5, having $\|r_k^i(t)\| > 0$ indicates that there exists a fault in the network other than $f_k(t)$. Additionally, since $r_j^i(t)$ is computed by a UIO decoupled from $f_j(t)$, if the only active fault is $f_j(t)$ we have $\|r_j^i(t)\| = 0$ and $\|r_k^i(t)\| > 0$ for all $k \neq j$. Motivated by this reasoning, we consider the following detection and isolation logic for fault $f_j(t)$ monitored by node i:

$$\begin{aligned} \left\|r_j^i(t)\right\| &< \Theta_j^i \\ \left\|r_k^i(t)\right\| &\geq \Theta_k^i \ , \forall k \neq j, \end{aligned} \tag{13}$$

where $\Theta_j^i > 0$ are isolation thresholds. These thresholds should be chosen according to trade-offs between sensitivity to faults, robustness to unmodeled dynamics and noise, misdetection rate, and false alarm rate, among others. Since choosing these thresholds is not within the scope of this paper, the reader is referred to [16] for further discussions.

Using Algorithm 1 a faulty node j can be detected and isolated by all the nodes in $\mathcal{N}_j$. However, all the other nodes in the network $i \notin \mathcal{N}_j$ can only detect the existence of a faulty node in the network, which occurs when $\left\|r_k^i(t)\right\| \geq \Theta_k^i \ \forall k \in \mathcal{N}_i$, while the identity of the faulty node is unknown to them. For the ease of notation we drop the superscript i from the variable names for the rest of this paper.

Algorithm 1 Distributed FDI of Faulty Nodes at Node i

for $k \in \mathcal{N}_i$ **do**
  Generate $r_k^i(t)$.
end for
if $\exists j : \left\|r_j^i(t)\right\| < \Theta_j^i \ \wedge \ \left\|r_k^i(t)\right\| \geq \Theta_k^i \ \forall k \in \mathcal{N}_i \neq j$ **then**
  Node j is faulty.
else if $\left\|r_k^i(t)\right\| \geq \Theta_k^i \ \forall k \in \mathcal{N}_i$ **then**
  There exists a faulty node $\ell \in \mathcal{V} \setminus \mathcal{N}_i$.
else if $\left\|r_k^i(t)\right\| < \Theta_k^i \ \forall k \in \mathcal{N}_i$ **then**
  There is no faulty node in the network.
end if

To solve the distributed FDI problem for faulty nodes using Algorithm 1, there needs to exist a bank of UIOs for each node $i \in \mathcal{V}$ satisfying the isolability condition in Proposition 2. For the case of faulty nodes, the problem of distributed FDI using UIOs can be stated as follows.

Problem 1. *Consider the networked system* (1) *and faulty nodes as in Definition 2. The answer to the following question is sought:*

1) *Consider the node j to be faulty, and let node i be a neighbor of j. Does there exist a UIO for node i that is decoupled from the faulty node j?*

The answer to Problem 1 has been provided in [22], where the authors prove the existence of matrices $F_k^i, T_k^i, K_k^i, H_k^i$ satisfying (10) for the system (8) with node faults and local measurements (3) for all $i \in \mathcal{V}$. In particular, the existence conditions of Proposition 1 reduce to having the graph $\mathcal{G}$ connected and $k \in \mathcal{N}_i$. Therefore we have the following assumption:

Assumption 1. *The network graph $\mathcal{G}$ is connected.*

B. Distributed FDI for Faulty Edges

In this section we extend the distributed FDI scheme to the case of faulty edges as in Definition 3. Similarly to the detection and isolation scheme outlined for node faults in Section III-A, faults on edges may also be detected and isolated using banks of UIOs. This section analyzes the existence of suitable UIOs that may be used to detects faulty edges. In particular, the following problem is addressed in this section.

Problem 2. *Consider the networked system* (1) *and faulty edges as in Definition 3. The answers to the following two questions are sought:*

1) *Consider the edge between nodes j and k to be faulty, and let node i be a neighbor of both j and k. Does there exist a UIO for node i that is decoupled from the faulty edge $\{j,k\}$?*
2) *Does there exist a UIO for node i that is decoupled from a faulty edge incident to node i?*

First we consider the problem of distributed detection and isolation of those faults that appear as corruptions in the communication or sensing links between pairs of neighbors characterized by Definition 3.1. Later the detection and isolation of edge parameter faults described in Definition 3.2 is tackled.

To address the problem of distributed detection and isolation of faulty edges, in addition to the bank of observers monitoring the fault in the neighbor nodes of a given node i to detect misbehaving nodes, we construct a bank of observers for those pairs of nodes neighboring to i that share the same edge. Hence at each node i, in addition to the observers for system models described by (8), observers for the following systems are constructed for all $\{j,k\} \in \mathcal{E}_i$:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bv(t) + E_{jk}f_{jk}(t) + E_{kj}f_{kj}(t) \\ y_i(t) &= C_i x(t) + D_{i,jk}f_{jk}(t) + D_{i,kj}f_{kj}(t) \end{aligned} \tag{14}$$

where $f_{jk}(t) = [f_{jk}^{\xi}(t) \ \ f_{jk}^{\zeta}(t)]^{\top}$, $E_{jk} = b_j[w_{jk} \ \ \mu w_{jk}]$, $D_{i,ij} = C_i[l_j \ \ b_j]$, and $D_{i,jk} = 0$ for $j \neq i$. Similarly as before, let $\hat{x}_{jk}(t)$ denote the estimate of the states for this system model and define the UIO decoupled from a faulty edge $\{j,k\}$ and the respective residual signal as follows.

Definition 6. *Consider the dynamical system* (14) *and the observer* (9)*. The observer is a* UIO decoupled from a faulty edge $\{j,k\}$ *if* $\lim_{t\to+\infty} \|x(t) - \hat{x}_{jk}^i(t)\| = 0$ *for any fault signals $f_{jk}(t)$ and $f_{kj}(t)$.*

Definition 7. *The signal $r_{jk}(t) \triangleq y_i(t) - C_i\hat{x}_{jk}(t)$ is a residual if $\|r_{jk}(t)\| = 0$ is equivalent to $\|f_{\bar{j}\bar{k}}(t)\| = \|f_{\bar{k}\bar{j}}(t)\| = 0$ for all $\{\bar{j},\bar{k}\} \neq \{j,k\} \in \mathcal{E}_i$.*

As seen in (14), the corrupted data sent along the faulty edge affects the dynamics of the node at the receiving end. In fact, comparing with the formulation in [21], [22], [25], such false data appears in the dynamics as two concurrent faulty nodes. However, note that the measurements $y_i(t)$ may also be affected by the edge fault. The following proposition establishes the existence of such observers for the system described above and addresses the first question posed in Problem 2.

Theorem 1. *Consider the networked system* (14) *with a sensing fault at the edge $\{j,k\}$ and $j,k \neq i$. In the sense of Definition 6, there exists a UIO decoupled from the faulty edge $\{j,k\}$ for node i if the graph $\mathcal{G}$ is connected and node i is a neighbor of both j and k.*

Proof. For node $i \in \mathcal{N}_j \cap \mathcal{N}_k$, the system dynamics and measurement equations are given by (14) with $E_{jk} =$

$b_j[w_{jk}\ \mu w_{jk}]$ and $D_{i,jk}=0$. Observing that the measurements at node i are not corrupted and defining $f^e_{jk}(t)=w_{jk}f^\xi_{jk}(t)+\mu w_{jk}f^\zeta_{jk}(t)$, the model can be rewritten as two simultaneous node faults:

$$\begin{aligned}\dot{x}(t)&=Ax(t)+E_{\{j,k\}}[f^e_{jk}(t)\ f^e_{kj}(t)]^\top\\ y_i(t)&=C_ix(t),\end{aligned}$$

with $E_{\{j,k\}}=[b_j\ b_k]$. Next we show that the UIO existence conditions in Proposition 1 are satisfied. It follows that the first rank condition in Proposition 1 holds because

$$\text{rank}(C_iE_{\{j,k\}})=\text{rank}(E^\top_{\{j,k\}}E_{\{j,k\}})=\text{rank}(E_{\{j,k\}}),$$

where $\text{rank}(C_iE_{\{j,k\}})=\text{rank}(E^\top_{\{j,k\}}E_{\{j,k\}})$ follows from the fact node i measures the states of nodes j and k that are affected by the fault.

As for the second rank condition in (12), it is the same as the case where two concurrent faults occur in the system, so the proof is similar to that of Theorem 1 in [22]. Consider the 1-hop neighborhood graph of node i, $\mathcal{G}_i$, with $\mathcal{V}_i=\{i\}\cup\mathcal{N}_i$ and $V_i=|\mathcal{V}_i|$. Denote $\tilde{\mathcal{G}}_i$ as the subgraph induced by the vertex set $\tilde{\mathcal{V}}_i=\mathcal{V}\backslash\mathcal{V}_i$, with $\tilde{V}_i=|\tilde{\mathcal{V}}_i|$. Without loss of generality, the nodes may be rearranged so that the Laplacian of $\mathcal{G}$ and $E_{\{j,k\}}$ can be written as

$$\mathcal{L}=\begin{bmatrix}\mathcal{L}_i & \ell_i\\ \ell_i^\top & \tilde{\mathcal{L}}_i\end{bmatrix},\quad E_{\{j,k\}}=\begin{bmatrix}0_{N\times 2}\\ l_{jk}\\ 0_{\tilde{V}_i\times 2}\end{bmatrix}$$

where $\ell_i\in\mathbb{R}^{V_i\times\tilde{V}_i}$ and the columns of $l_{jk}\in\mathbb{R}^{V_i\times 2}$ are the columns of I_{V_i} corresponding to nodes j and k. The second rank condition in (12) becomes

$$\text{rank}\underbrace{\begin{bmatrix}sI_{V_i} & 0_{V_i\times\tilde{V}_i} & -I_{V_i} & 0_{V_i\times\tilde{V}_i} & 0_{V_i\times 2}\\ 0_{\tilde{V}_i\times V_i} & sI_{\tilde{V}_i} & 0_{\tilde{V}_i\times V_i} & -I_{\tilde{V}_i} & 0_{\tilde{V}_i\times 2}\\ \mathcal{L}_i & \ell_i & \alpha_1(s) & \mu\ell_i & l_{jk}\\ \ell_i^\top & \tilde{\mathcal{L}}_i & \mu\ell_i^\top & \alpha_2(s) & 0_{\tilde{V}_i\times 2}\\ I_{V_i} & 0_{V_i\times\tilde{V}_i} & 0_{V_i\times V_i} & 0_{V_i\times\tilde{V}_i} & 0_{V_i\times 2}\\ 0_{V_i\times V_i} & 0_{V_i\times\tilde{V}_i} & I_{V_i} & 0_{V_i\times\tilde{V}_i} & 0_{V_i\times 2}\end{bmatrix}}_{P}=2N+2,$$

where $\alpha_1(s)=sI_{V_i}+\mu\mathcal{L}_i+\bar{K}_i$ and $\alpha_2(s)=sI_{\tilde{V}_i}+\mu\tilde{\mathcal{L}}_i+\tilde{K}_i$.

Observing that the first and third column blocks are linearly independent of the rest and applying some row and column operations we have

$$\text{rank}(P)=\text{rank}\begin{bmatrix}-\frac{1}{\mu}I_{\tilde{V}_i} & -(1+\mu s)I_{\tilde{V}_i} & 0_{\tilde{V}_i\times 2}\\ \ell_i & 0_{V_i\times\tilde{V}_i} & l_{jk}\\ 0_{\tilde{V}_i\times\tilde{V}_i} & -\alpha(s) & 0_{\tilde{V}_i\times 2}\end{bmatrix}+2V_i,$$

with $\alpha(s)=\mu s^2I_{\tilde{V}_i}+\mu s(\tilde{\mathcal{L}}_i+\tilde{K}_i)+\tilde{\mathcal{L}}_i$. It follows from [29] that $\tilde{\mathcal{L}}_i$ is positive definite if $\mathcal{G}$ is connected. Since $\mu>0$ and $\tilde{K}_i$ are positive definite, we conclude that $\alpha(s)$ is invertible for $s\in\mathbb{C}$ with non-negative real part. Therefore the first and second column blocks are independent of each other and the third column block, which concludes the proof. □

Moreover we have the following result stating that, for any node i, an observer decoupled from a faulty edge incident to i cannot be constructed. It addresses the second question posed in Problem 2.

Proposition 3. *Consider the networked system* (14) *with a sensing fault at the edge* $\{i,j\}$*. In the sense of Definition 6, there does not exist a UIO decoupled from the faulty edge* $\{i,j\}$ *for node* i*.*

Proof. Consider a faulty edge $\{i,j\}$ incident to node i with a sensing fault. Recalling (14), the system dynamics and measurement equations can be rewritten as

$$\begin{aligned}\dot{x}(t)&=Ax(t)+Bv(t)+E_{\{i,j\}}f_{\{i,j\}}(t)\\ y_i(t)&=C_ix(t)+D_{i,\{i,j\}}f_{\{i,j\}}(t)\end{aligned}$$

where $f_{\{i,j\}}(t)=[f_{ij}^\top(t)\ f_{ji}^\top(t)]^\top$, $E_{\{i,j\}}=[E_{ij}\ E_{ji}]$ and $D_{i,\{i,j\}}=[D_{i,ij}\ 0]$. From [16] we recall that the following rank condition should hold for the existence of UIOs:

$$\text{rank}\begin{bmatrix}D_{i,\{i,j\}} & C_iE_{\{i,j\}}\\ 0 & D_{i,\{i,j\}}\end{bmatrix}=\text{rank}(D_{i,\{i,j\}})+\text{rank}\begin{bmatrix}E_{\{i,j\}}\\ D_{i,\{i,j\}}\end{bmatrix},$$

where the second term equals 5. Given $C_iE_{\{i,j\}}$ and $D_{i,\{i,j\}}$, the first term of the latter rank condition can be written as

$$\text{rank}\begin{bmatrix}C_il_j & C_ib_j & C_ib_iw_{ij} & C_ib_i\mu w_{ij}\\ 0 & 0 & C_il_j & C_ib_j\end{bmatrix}\leq 4,$$

since each column-block is a column vector. Since the rank condition is not fulfilled, there does not exist a UIO for this system. □

Although in the case of bidirectional sensing faults in edges there is no UIO for the nodes to which the faulty edge is incident to, the following result shows that this is not the case for unidirectional faults, i.e., for the case where either $f_{ij}(t)$ or $f_{ji}(t)$ is identically zero. We formalize this case in what follows.

Proposition 4. *Consider the networked system* (14) *with a sensing fault at the edge* $\{i,j\}$*. In the sense of Definition 6, if the graph* $\mathcal{G}$ *is connected, for node* i *there exist a UIO decoupled from*

1) *The sensing fault from node* j *to node* i*,* $f_{ij}(t)$*, when* $f_{ji}(t)\equiv 0$*.*
2) *The sensing fault from node* i *to node* j*,* $f_{ji}(t)$*, when* $f_{ij}(t)\equiv 0$*.*

Proof. In the first case, the dynamical system with respect to node i and the faulty edge $\{i,j\}$ is described by (14) with $E_{ij}=b_i[w_{ij}\ \mu w_{ij}]$, $E_{ji}=0$, $D_{ij}=C_i[l_j\ b_j]$, and $D_{ji}=0$. Now consider that the measurements corresponding to node j have been removed, yielding the following system

$$\begin{aligned}\dot{x}(t)&=Ax(t)+Bv(t)+E_{ij}f_{ij}(t),\\ \tilde{y}_i(t)&=\tilde{C}_ix(t),\end{aligned}$$

which corresponds to the model of a single node fault at node i and measurements from $\mathcal{V}_i^1\backslash\{j\}$. From [22], it then follows that a UIO exists for this system.

In the second case, the dynamical system with respect to node i is described by

$$\begin{aligned}\dot{x}(t)&=Ax(t)+Bv(t)+E_{ji}f_{ji}(t)\\ y_i(t)&=C_ix(t)\end{aligned}$$

which also corresponds to a single node fault at node j and, similarly to the previous case, the corresponding UIO exists. □

In the following we consider faulty edges with parameter faults, as described in Definition 3.2. For detecting and isolating these faults at each node i, in addition to the observers for system models described by (8), observers for the following systems are constructed at each node i for all $\{j,k\} \in \mathcal{E}_i$:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bv(t) + E_{jk} f_{jk}(t) \\ y_i(t) &= C_i x(t) \end{aligned} \tag{15}$$

where $E_{jk} = b_j - b_k$ and $f_{jk}(t) = \delta_{jk}(t) f^w_{jk}(t)$. The existence of UIO's for (15) is a consequence of the results establishing the existence of UIO's for faulty nodes and will not be stated here for brevity.

Under the assumption that a single fault occurs at any given time, the following algorithm may be implemented at each node to simultaneously detect and isolate faulty nodes and edges.

Algorithm 2 Distributed FDI of Faulty Nodes and Edges at Node i

for $j \in \mathcal{N}_i$ **do**
  Generate $r_j(t)$.
end for
for $\{j,k\} \in \mathcal{E}_i$ **do**
  Generate $r_{jk}(t)$.
end for
if $\exists k: \ \|r_k(t)\| < \Theta_k$ **and** $\|r_j(t)\| \geq \Theta_j, \ \forall j \in \mathcal{N}_i \neq k$ **then**
  Node k is faulty.
end if
if $\exists \{\bar{j},\bar{k}\}: \ \left\|r_{\bar{j}\bar{k}}(t)\right\| < \Theta_{\{\bar{j},\bar{k}\}}$ **and** $\|r_j(t)\| \geq \Theta_j, \ \forall j \in \mathcal{N}_i \neq k$ **and** $\|r_{jk}(t)\| \geq \Theta_{\{j,k\}}, \ \forall \{j,k\} \in \mathcal{E}_i \neq \{\bar{j},\bar{k}\}$ **then**
  Edge $\{\bar{j},\bar{k}\}$ is faulty.
end if
if $\|r_j(t)\| \geq \Theta_j \ \forall j \in \mathcal{N}_i$ **and** $\|r_{jk}(t)\| \geq \Theta_{\{j,k\}} \ \forall \{j,k\} \in \mathcal{E}_i$ **then**
  There exists a faulty node or edge in $\mathcal{G} \backslash \mathcal{G}_i$.
end if
if $\|r_j(t)\| < \Theta_j \ \forall j \in \mathcal{N}_i$ **and** $\|r_{jk}(t)\| < \Theta_{\{j,k\}} \ \forall \{j,k\} \in \mathcal{E}_i$ **then**
  There is no faulty node or edge in the network.
end if

IV. Distributed FDI in the Presence of Imprecise Network Model

As described earlier, to construct a bank of observers achieving distributed FDI given the local measurements (3), the knowledge of the system matrix A is needed. In this section we study the case where, after having designed observers under a known network model and interconnection graph, some edges and nodes are removed. The edge and node removal may correspond to either unexpected changes in the system, or the removal of faulty edges and nodes. In both scenarios, it is desirable to maintain the detection and isolation capabilities of the distributed FDI scheme despite the model changes. Later in this section we show that a distributed FDI scheme does not require the full knowledge of the network. Now we are ready to pose the following problem.

Problem 3. *Consider a network and a bank of observers as described in Section III. Suppose the network loses l edges. What are the necessary and sufficient conditions ensuring that node i can detect faults in the network using the bank of observers and Algorithm 2?*

Note that removing a node corresponds to removing all the edges incident to it, thus the case of node removal is covered by the previous problem.

A. Distributed FDI with global model

We first address Problem 3 when the global model (8) is used to design the UIOs. Consider the case where we design a bank of UIO's to estimate the states of the neighbors of node i and recall that we have the following observer error and residual dynamics

$$\begin{aligned} \dot{e}_k(t) &= F_k e_k(t) - T_k \sum_{m \in \mathcal{N}_i \backslash \{k\}} E_m f_m(t) \\ r_k(t) &= C_i e_k(t). \end{aligned} \tag{16}$$

Introduce $\mathcal{E}_{loss} \subseteq \mathcal{E}$ as the subset of edges removed from the network. Recalling the system dynamics (8), under edge removal the new system and output matrices A_ℓ and $C_{i\ell}$, respectively, are given by

$$\begin{aligned} A_\ell &= A + \Delta A, \\ C_{i\ell} &= C_i + \Delta C_i. \end{aligned} \tag{17}$$

The matrices ΔA and ΔC_i are perturbation matrices corresponding to the lost edges. More precisely, $\Delta A = \begin{bmatrix} 0_N & 0_N \\ \mathcal{L}_{loss} & \mu \mathcal{L}_{loss} \end{bmatrix}$, where $\mathcal{L}_{loss}$ is the Laplacian matrix corresponding to the graph $\mathcal{G}_{loss}(\mathcal{V}, \mathcal{E}_{loss})$. Moreover, all the entries of ΔC_i are zero except those entries that correspond to a neighbor of i whose shared edge with i is in $\mathcal{E}_{loss}$, which are all equal to -1. We have the following assumption.

Assumption 2. *The network remains connected after losing the edges $\mathcal{E}_{loss}$.*

Using the existing parameters of the UIO (computed under the assumption of no edge loss), the error dynamics are characterized by

$$\begin{aligned} \dot{e}_k(t) =& F_k e_k(t) + \Delta A x(t) + H_k C_i \Delta A x(t) + H_k \Delta C_i \Delta A x(t) \\ & - K_k \Delta C_i x(t) - T_k \sum_{m \in \mathcal{N}_i \backslash \{k\}} E_m f_m(t). \end{aligned} \tag{18}$$

If the removed links had not been connecting i to any of its neighbors, we have $\Delta C_i = 0$. It is easy to check that then the error dynamics become

$$\dot{e}_k(t) = F_k e_k(t) + (I + H_k C_i) \Delta A x(t) - T_k \sum_{m \in \mathcal{N}_i \backslash \{k\}} E_m f_m(t). \tag{19}$$

The error dynamics described by (19), in the presence of no faults for $m \in \mathcal{V} \setminus \{k\}$, $f_m(t) \equiv 0$, are

$$\dot{e}_k(t) = F_k e_k(t) + (I + H_k C_i)\Delta A x(t). \tag{20}$$

Assume for the moment that the known input $v(t)$ is zero. Recall from Remark 1 that, if the network is connected, $x(t)$ converges exponentially to $[\bar{\xi}\ \bar{\zeta}]^\top \otimes \mathbf{1}_{2N}$ when there is no fault. Given the structure of ΔA and recalling that $\mathcal{L}\mathbf{1}_N = 0$ for any Laplacian matrix $\mathcal{L} \in \mathbb{R}^{N\times N}$, it follows that $\Delta A x(t)$ goes exponentially fast to zero when there is no fault in the network. Therefore, since F_k is Hurwitz, the error dynamics described by (20) are stable. Consequently $r_k(t) = C_i e_k(t)$ goes to zero when there is no fault in the system, although the UIO parameters are designed for a different interconnection network. However, if $v(t) \neq 0$ does not drive the system to consensus, i.e. $\|x_i(t) - x_j(t)\|$ does not go to zero as t goes to infinity, then $\Delta A x(t)$ does not generically converge to zero when there is no fault, and neither does the residual $r_k(t)$.

On the other hand, if any of the removed edges had been connecting i to one of its neighbors, the error dynamics may not even converge to zero when there is no fault. In particular, suppose there are no faults and that the system has reached an equilibrium so that $\Delta A x(t) = 0$, yielding the error dynamics

$$\dot{e}_k(t) = F_k e_k(t) - K_k \Delta C_i x(t). \tag{21}$$

Since in general $K_k \Delta C_i x(t)$ is not identically zero at the equilibrium, we conclude that the error does not converge to zero and thus $r_k(t)$ is not a suitable residual, as it violates Definition 5. Hence, the bank of observers should be redesigned taking into account the updated network model. Formally, we have the following result that addresses Problem 3.

Theorem 2. *Consider a monitoring node i in an arbitrary connected network described by* (1) *and a bank of UIO's for this network. Using Algorithm 1 and the existing bank of observers, node i can detect the presence of a faulty node after the loss of ℓ edges if and only if all the following conditions are satisfied: (1) the network remains connected, (2) $v(t)$ is such that $\|x_i(t) - x_j(t)\| \to 0$ as $t \to \infty$, i.e. it drives the system to consensus, and (3) $\mathcal{N}_i$ is the same as in the original network.*

Proof. Consider the original graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ and let $k \in \mathcal{N}_i$. Suppose ℓ edges in set $\tilde{\mathcal{E}}$ are lost and the corresponding subgraph to these edges is denoted by $\tilde{\mathcal{G}}(\mathcal{V}, \tilde{\mathcal{E}})$. Since i cannot detect faults in network components it is not connected to, a necessary condition is that the subgraph $\tilde{\mathcal{G}}$ remains connected. Another necessary condition is that $v(t)$ drives the system to consensus, thus ensuring that $\Delta A x(t)$ does converge to zero. Additionally, having $\Delta C_i = 0$, or equivalently $k \in \tilde{\mathcal{N}}_i$ for all $k \in \mathcal{N}_i$, is also a necessary condition. Otherwise, in general the residuals do not converge to zero.

Now suppose all the necessary conditions hold. When there is no fault in the network, $e_k(t)$ goes to zero and as a result $\|r_k(t)\|$ goes to zero as well. For the faulty case, $\|r_k(t)\|$ will generically not converge to zero for $k \in \mathcal{N}_i$. Hence, using Algorithm 1 one can detect if there is a fault in the network or not. □

Note that the faulty node cannot be isolated using the condition given by (13) when the network model is imprecise. Moreover, detection is also not feasible when the system is not driven to consensus by $v(t)$.

These limitations follow from the fact that $\Delta A x(t)$ does not go to zero because, in general, $x(t)$ does not reach consensus under the fault $f_k(t)$. Thus, the error of the UIO monitoring the neighbor node k converges to a ball around zero with a nonzero radius. Hence, none of the residuals goes to zero so (13) cannot be used to isolate the faulty node.

A possible way to overcome such limitations is to use additional measurements from outside each node's neighborhood and design the bank of UIOs using local models of the system that are not not affected by changes in other parts of the network. In particular, we consider the following problem.

Problem 4. *For a given node i, consider a subgraph of the network $\tilde{\mathcal{G}}_i$ containing the 1-hop neighborhood graph $\mathcal{G}_i$. Let any state measurement within $\tilde{G}_i$ be available to node i. The following questions are considered:*

- *For which subgraphs can node i design a bank of UIOs and implement Algorithm 1 to detect and isolate faults in any of its neighbors?*
- *Given the set of subgraphs for which a UIO-based FDI scheme exists, which subgraph $\tilde{\mathcal{G}}_i$ minimizes the number of edges in $\tilde{\mathcal{G}}_i$ and required state measurements?*

In what follows we propose a method to address the problem of isolating the faulty nodes and edges in the network, and tackle Problem 4.

B. Distributed FDI with local models

Consider a fault-free network $\mathcal{G}(\mathcal{V}, \mathcal{E})$ with the system dynamics $\dot{x}(t) = Ax(t) + Bv(t)$. Define $\hat{\mathcal{G}}_i$ as a subgraph containing the proximity subgraph of node i, $\mathcal{P}_i \subseteq \hat{\mathcal{G}}_i \subseteq \mathcal{G}(\mathcal{V}, \mathcal{E})$. Let $\mathcal{B}(\hat{\mathcal{V}}_i) \subseteq \hat{\mathcal{V}}_i$ be the boundary vertex set such that $\ell \in \mathcal{B}(\hat{\mathcal{V}}_i)$, if $\{\ell, \bar{\ell}\} \in \mathcal{E}$ and $\bar{\ell} \notin \hat{\mathcal{V}}_i$.

The dynamics of the subsystem associated with $\hat{\mathcal{G}}_i$ are

$$\dot{\phi}^i(t) = A^i_{\hat{\mathcal{G}}} \phi^i(t) + \psi^i(t) + B^i_{\hat{\mathcal{G}}} v^i_{\hat{\mathcal{G}}}(t), \tag{22}$$

where $\phi_i = [\xi_i\ \xi_{i_1}\ \dots\ \xi_{i_{|\hat{\mathcal{V}}_i|}}\ \zeta_i\ \zeta_{i_1}\ \dots\ \zeta_{i_{|\hat{\mathcal{V}}_i|}}]$, $i_m \in \hat{\mathcal{V}}_i$. Particularly i_1 to $i_{|\mathcal{N}_i|}$ are associated with the nodes in $\mathcal{N}_i$. Moreover, $A^i_{\hat{\mathcal{G}}}$ is the matrix associated with the network with $\hat{\mathcal{G}}_i$ as its graph, $\psi^i(t)$ is an unknown vector with zero entries except for the entries corresponding to nodes $j \in \mathcal{B}(\hat{\mathcal{V}}_i)$ that represents the interaction of the rest of the network with the subnetwork of interest. Additionally, $v^i_{\hat{\mathcal{G}}}(t)$ is an input vector in this subnetwork known to i, and $B^i_{\hat{\mathcal{G}}}$ is the input matrix associated with these inputs. We have the following straightforward result for $\psi^i(t)$.

Proposition 5. *In the network induced by the proximity graph of node i as described by* (22)*, $\psi^i(t)$ goes to zero exponentially fast for $v(t) \equiv 0$.*

Proof. The proof is a direct consequence of the exponential stability of (1) to the consensus equilibrium and the distributed control law (4). □

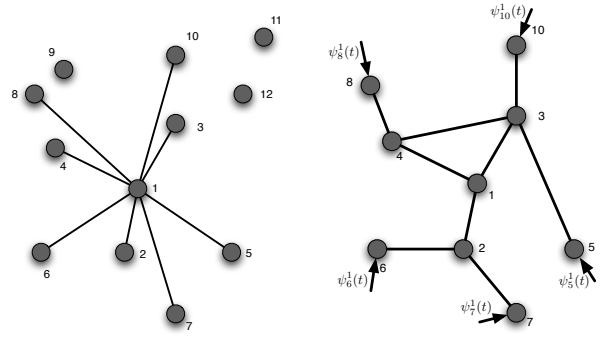
The bank of UIOs at i can be designed for the subnetwork with $\hat{\mathcal{G}}_i$ as its graph and dynamics described by (22). An example of such a subnetwork for the network of Fig. 2 when $\hat{\mathcal{G}}_i = \mathcal{P}_i$ is given in Fig. 3 (b).

In the case where there is no fault in the network and $v(t) \equiv 0$, the unknown parts of the real network enter the equation dynamics as exponentially decaying signals. As before, in this case the detection of a fault can be determined using the bank of UIOs for $\hat{\mathcal{G}}_i$. Moreover, isolation can be achieved by choosing an appropriate threshold value.

However, the selection of the threshold might be cumbersome, and it requires a knowledge of the magnitude of the fault. In what comes next we propose a method to achieve distributed FDI using only the full knowledge of the subgraph graph $\hat{\mathcal{G}}_i$, without resorting to complicated ways of choosing the threshold value and allowing $v(t) \not\equiv 0$. Given $\hat{\mathcal{G}}_i$, let $\mathcal{S}_i(\hat{\mathcal{V}}_i) \subseteq \hat{\mathcal{V}}_i$ be the set of the nodes for which node i measures states. We make the following assumption that will be valid until the end of this section.

Assumption 3. *For each node $i \in \mathcal{V}$ and the corresponding subgraph $\hat{\mathcal{G}}_i(\hat{\mathcal{V}}_i, \hat{\mathcal{E}}_i) \subseteq \mathcal{G}(\mathcal{V}, \mathcal{E})$ containing the proximity graph $\mathcal{P}_i$, the state measurements of nodes in $\mathcal{S}_i(\hat{\mathcal{V}}_i) = \{i\} \cup \mathcal{N}_i \cup \mathcal{B}(\hat{\mathcal{V}}_i)$ are available to node i.*

An example for the measurement graph of node i is given in Fig. 3(a). As before, to achieve the fault detection and isolation



(a) (b)

Fig. 3. (a) An example of a measurement graph of node i in the network of Fig. 2 under Assumption 3. (b) The subnetwork used for designing a bank of UIO's at node 1 of the network depicted in Fig. 2.

task each node i considers $|\mathcal{N}_i|$ models of the form:

$$\dot{\phi}^i(t) = A^i_{\hat{\mathcal{G}}}\phi^i(t) + \psi^i(t) + B^i_{\hat{\mathcal{G}}}v^i_{\hat{\mathcal{G}}}(t) + E^i_k f_k(t) \tag{23}$$

where E^i_k is a vector of zeros except for the entry corresponding to node $k \in \mathcal{N}_i$, which is equal to one. We rewrite (23) as

$$\dot{\phi}^i(t) = A^i_{\hat{\mathcal{G}}}\phi^i(t) + B^i_{\hat{\mathcal{G}}}v^i_{\hat{\mathcal{G}}}(t) + \begin{bmatrix} E^i & E^i_k \end{bmatrix} \begin{bmatrix} \psi^i(t) \\ f_k(t) \end{bmatrix}, \tag{24}$$

with $E^i = [E^i_{m_1} \ \dots \ E^i_{m_{|\mathcal{B}(\hat{\mathcal{V}}_i)|}}]$, where $E^i_{m_l}$, $m_l \in \mathcal{B}(\hat{\mathcal{V}}_i)$, is a vector of zeros except for the entry corresponding to node $m_l \in \mathcal{B}(\hat{\mathcal{V}}_i)$ that is equal to one. For each of these models, a UIO that is decoupled from the unknown input $\begin{bmatrix} E^i & E^i_k \end{bmatrix} \begin{bmatrix} \psi^i(t) \\ f_k(t) \end{bmatrix}$ is designed.

Lemma 1. *Consider the distributed control system with a fault in node $j \in \mathcal{N}_i$ given by* (22) *and measurements satisfying Assumption 3. In the sense of Definition 4, there exists a UIO for node i that is decoupled from the faulty node j and the subgraph $\mathcal{V}\backslash\hat{\mathcal{V}}_i$.*

Proof. Recall the UIO existence condition in Proposition 1. From Assumption 3, node i measures its own states, as well as the states of nodes $j \in \mathcal{B}(\hat{\mathcal{V}}_i)$ and $j \in \mathcal{N}_i$, which are the ones affected by the unknown input $\psi^i(t)$ and the fault $f_j(t)$, respectively. Therefore it follows that $\text{rank}(C_i E^i) = \text{rank}(E^{i\top} E^i)$ and $\text{rank}(C_i E^i_k) = \text{rank}(E^{i\top}_k E^i_k)$, thus the first rank condition holds.

As for the second rank condition in (12), consider the subgraph $\tilde{\mathcal{G}}_i$ induced by the vertex set $\tilde{\mathcal{V}}_i = \mathcal{B}(\hat{\mathcal{V}}_i)$ with $\tilde{V}_i = |\tilde{\mathcal{V}}_i|$. Denote $\bar{\mathcal{G}}_i$ as the subgraph induced by the vertex set $\bar{\mathcal{V}}_i = \hat{\mathcal{V}}_i\backslash\tilde{\mathcal{V}}_i$, with $\bar{V}_i = |\bar{\mathcal{V}}_i|$ and note that $\hat{V}_i \triangleq |\hat{\mathcal{V}}_i| = \tilde{V}_i + \bar{V}_i$. Without loss of generality, the nodes may be rearranged so that the Laplacian of $\hat{\mathcal{G}}_i$, E^i_k, E^i, and C_i can be written as

$$\hat{\mathcal{L}} = \begin{bmatrix} \bar{\mathcal{L}}_i & \bar{\ell}_i \\ \bar{\ell}_i^\top & \tilde{\mathcal{L}}_i \end{bmatrix}, \ E^i_k = \begin{bmatrix} 0_{\hat{V}_i \times 1} \\ l_k \\ 0_{\tilde{V}_i \times 1} \end{bmatrix}, \ E^i = \begin{bmatrix} 0_{\hat{V}_i \times \tilde{V}_i} \\ 0_{\bar{V}_i \times \tilde{V}_i} \\ I_{\tilde{V}_i} \end{bmatrix},$$

$$C_i = \begin{bmatrix} \bar{C}_i & 0_{V_i \times \tilde{V}_i} & 0_{V_i \times \bar{V}_i} & 0_{V_i \times \tilde{V}_i} \\ 0_{\tilde{V}_i \times \bar{V}_i} & I_{\tilde{V}_i} & 0_{\bar{V}_i \times \bar{V}_i} & 0_{\bar{V}_i \times \tilde{V}_i} \\ 0_{V_i \times \bar{V}_i} & 0_{V_i \times \tilde{V}_i} & \bar{C}_i & 0_{V_i \times \tilde{V}_i} \\ 0_{\tilde{V}_i \times \bar{V}_i} & 0_{\tilde{V}_i \times \tilde{V}_i} & 0_{\bar{V}_i \times \bar{V}_i} & I_{\tilde{V}_i} \end{bmatrix},$$

where $\bar{\ell}_i \in \mathbb{R}^{\bar{V}_i \times \tilde{V}_i}$, $l_k \in \mathbb{R}^{\bar{V}_i \times 1}$, and $\bar{C}_i \in \mathbb{R}^{|\mathcal{V}^1_i| \times \bar{V}_i}$ being a full row rank matrix where each of the rows have all zero entries except for one entry at the j-th position that corresponds to those nodes that are in $\mathcal{V}^1_i = \mathcal{N}_i \cup \{i\}$. Following a similar reasoning as in Theorem 1, one can verify that the second rank condition in (12) also holds. □

Such UIO scheme can clearly be implemented for any subgraph $\hat{\mathcal{G}}_i$ containing the proximity graph $\mathcal{P}_i$. Applying Algorithm 1 or Algorithm 2 for the residuals obtained from these UIOs, with $\mathcal{G}$ replaced with $\hat{\mathcal{G}}_i$, addresses the first part of Problem 4. Hence, node i can detect and isolate a fault in node $j \in \mathcal{N}_i$ using only local models and measurements, as stated in the following result.

Theorem 3. *Consider a monitoring node i in a connected network satisfying Assumption 2 and a bank of UIO's calculated for the local subsystem* (24)*. Using Algorithm 1 and the bank of observers, node i can detect and isolate a faulty node in its neighborhood.*

Proof. The proof follows from Lemma 1 and Theorem 2. □

V. Complexity Reduction of Distributed FDI

So far we have proposed the solutions to both Problems 3 and 4. In Section IV we first showed that it is possible to detect the presence of a faulty node in the network distributedly, i.e., address Problem 3, at each node i, if i knows the exact

model of its one-hop neighborhood and measuring the states of its neighbors. Then we introduced a method to address the first part of Problem 4 that not only eliminates the need to have an exact network model beyond a subgraph containing the proximity graph of a given node for that node to detect and isolate faults in its one-hop neighborhood, but it also reduces the size of the observers. However, such result is derived under the assumption that the node has access to all the measurements of the states of its two-hop neighbors. In this section we show that the knowledge of the proximity graph is in fact the least amount of knowledge required to achieve distributed FDI when equal costs are associated with each necessary state measurement and network component that needs to be known, thus addressing the second part of Problem 4. Later, the complexity of the overall distributed FDI scheme is minimized by reducing the number of monitoring nodes while still ensuring that every node in the network is monitored.

A. Local models and additional measurements

Suppose node i has the local model (24) for a given subgraph $\hat{\mathcal{G}}_i(\hat{\mathcal{V}}_i, \hat{\mathcal{E}}_i)$. Consider the case where equal costs are associated with each node ℓ in $\mathcal{B}(\hat{\mathcal{V}}_i)$, and with each of the edges that are known exactly, i.e., each $\{j,k\} \in \hat{\mathcal{E}}_i$. In other words, a cost is associated with any piece of information available to a node i; be it extra measurements or information about the existence of an edge between two nodes. This cost is minimized by solving the following optimization problem:

$$\min_{\mathcal{P}_i \subseteq \hat{\mathcal{G}}_i \subseteq \mathcal{G}} |\mathcal{S}_i(\hat{\mathcal{V}}_i)| + |\hat{\mathcal{E}}_i|. \tag{25}$$

We conclude this section by introducing the following result that shows that knowing $\mathcal{P}_i$ exactly is optimal, in the sense that it minimizes (25).

Theorem 4. *Consider a monitoring node i in an arbitrary connected network and a bank of UIO's calculated for the local subsystem $\hat{\mathcal{G}}_i$. Setting $\hat{\mathcal{G}}_i = \mathcal{P}_i$ simultaneously minimizes the number of state measurements $|\mathcal{S}_i|$ and the number of known network connections $|\hat{\mathcal{E}}_i|$ needed to design the bank of UIO's.*

Proof. Recall from Assumption 3 that $\mathcal{S}_i(\hat{\mathcal{V}}_i) = \{i\} \cup \mathcal{N}_i \cup \mathcal{B}(\hat{\mathcal{V}}_i)$. From Lemma 1 we know that any $\hat{\mathcal{G}}_i$ should be such that $\mathcal{P}_i \subseteq \hat{\mathcal{G}}_i$. To obtain a contradiction, assume that there is a $\mathcal{G}_i^\star(\mathcal{V}_i^\star, \mathcal{E}_i^\star)$ such that $\mathcal{P}_i$ is a strict subset of $\mathcal{G}_i^\star(\mathcal{V}_i^\star, \mathcal{E}_i^\star)$ that results in a smaller value for the objective function in (25). We can obtain it by adding vertices that are in $\mathcal{V}_i^\star \setminus \mathcal{V}_{P_i}$ one by one to $\mathcal{P}_i$. If we introduce a single vertex ℓ_1 to $\mathcal{P}_i$, then it is necessary that all the $\bar{\eta}$ edges $\{\ell_1, j\}$ such that $j \in \mathcal{V}_{\mathcal{P}_i}$ are exactly known, in addition to all the η edges incident to the vertices in $\mathcal{N}_i^2$. Call this new graph obtained from the addition of ℓ_1 and the aforementioned edges $\mathcal{G}_i^{+\ell_1}(\mathcal{V}_i^{+\ell_1}, \mathcal{E}_i^{+\ell_1})$. Then we have

$$\begin{aligned} |\mathcal{B}(\mathcal{V}_i^{+\ell_1})| + |\mathcal{E}_i^{+\ell_1}| &= |\mathcal{B}(\mathcal{V}_{P_i})| - \eta + 1 + |\mathcal{E}_{P_i}| + \eta + \bar{\eta} \\ &= |\mathcal{B}(\mathcal{V}_{P_i})| + 1 + |\mathcal{E}_{P_i}| + \bar{\eta}. \end{aligned} \tag{26}$$

Even for the case where there are no edges in the network connecting the nodes in $\mathcal{N}_i^2$, i.e., $\bar{\eta} = 0$, the cost function is increased by at least one. Repeating this argument for addition of any other vertex $\ell_j \in \mathcal{V}_i^\star \setminus \mathcal{V}_{P_i}$, one can deduce that the cost function does not decrease. Hence, there exists no $\mathcal{G}_i^\star$, such that $\mathcal{P}_i \not\subseteq \mathcal{G}_i^\star$, that minimizes the cost function given in (25). □

Theorem 4 provides the optimal subgraph $\hat{\mathcal{G}}_i$ that minimizes the amount of model knowledge and number of measurements where they are equally valued. However, if the cost of having measurements from a node is equal to $c_m \geq 0$ and the cost of knowing the existence of an edge is equal to $c_e \geq 0$, and $c_m \neq c_e$, (25) becomes

$$\min_{\mathcal{P}_i \subseteq \hat{\mathcal{G}}_i \subseteq \mathcal{G}} c_m |\mathcal{S}_i(\hat{\mathcal{V}}_i)| + c_e |\hat{\mathcal{E}}_i|. \tag{27}$$

One can construct simple examples with $c_m \neq c_e$ where taking $\hat{\mathcal{G}}_i = \mathcal{P}_i$ does not necessarily minimize the cost function proposed in (27).

B. Reducing the number of monitoring nodes

It is not necessary for all the nodes in a network to monitor their neighbors and it is possible to decrease the number of monitoring nodes in the network while guaranteeing that each node in the network is being monitored by at least another node and calculating UIO's for only these nodes.

Assuming that each node monitors only its neighbors, we say that a FDI system in node i covers the set of nodes $\mathcal{N}_i$. Therefore, the objective is to select a minimum number of observer nodes that cover all the nodes in the network, i.e.,

$$\begin{aligned} \min_{S_o \subseteq \mathcal{V}} \quad & |S_o| \\ s.t. \quad & \bigcup_{i \in S_o} \mathcal{N}_i = \mathcal{V}, \end{aligned} \tag{28}$$

where S_o is the set of observer nodes.

As it can be seen, this is actually a set cover problem where we wish to determine a *minimum total dominating set, i.e.,* a set with minimum cardinality such that all nodes in the graph have at least one neighbor in that set. This is a well studied problem, having been classified as an NP-hard problem and two algorithms to solve this problem can be found in [30].

Although the number of observers obtained by using $\mathcal{N}_i$ as the set of nodes covered by node i is not minimum, this method has one interesting property: all nodes in S_o are monitored by at least one neighbor. This means that even if an observer node is attacked, there is another observer node in the network that can detect it. Obviously, this decreases the vulnerability to faults in the monitoring nodes.

Other interesting properties may also be imposed by modifying the constraints in (28), such as having S_o to be connected, which is related to the *minimum connected dominating set* problem.

Another way of minimizing the computational burden of the proposed method is to find a set of nodes that monitors all the nodes in the network with the minimum number of measurements, i.e., solving (28) with the cost function $|S_o|$ replaced with $\sum_{i \in S_o} \deg(i)$. This problem can be solved first by

finding all the dominating sets in the network and choosing the set that minimizes the cost function.

VI. Numerical Examples

In this section we illustrate the solution proposed in the paper on a power network example. The simulations were carried out using the IEEE 118 bus network example available with the MATPOWER toolbox [31]. A diagram of the power network is depicted in Fig. 4.

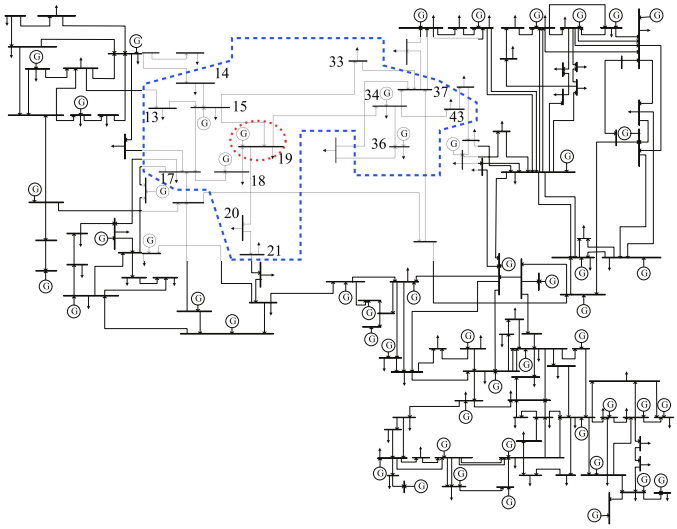


Fig. 4. Diagram of the IEEE 118 bus power network. The monitoring node 19 is encircled with a red dotted line, while its 2-hop neighborhood is delimited by the blue dashed line.

We considered the classical linearized synchronous machine model [28] for each node of the power network, leading to the global network dynamics as in (1) with

$$A = \begin{bmatrix} 0_N & I_N \\ -\bar{M}\mathcal{L} & -\bar{M}\bar{D} \end{bmatrix}, \quad B = \begin{bmatrix} 0_N & \bar{M} \end{bmatrix}^\top,$$
$$\bar{M} = \text{diag}\left(\frac{1}{m_1}, \cdots, \frac{1}{m_N}\right), \quad \bar{D} = \text{diag}\left(d_1, \cdots, d_N\right),$$

where $m_i > 0$ and $d_i > 0$ are the inertia and damping coefficients of node i, respectively, and $N = 118$ is the number of buses. Since these coefficients were not available in the example data files, they were randomly generated so that the load buses had considerably lower values than the generator buses, namely $m_g \approx 10^3 m_l$ and $d_g \approx 10^3 d_l$.

A. Faulty node detection using local model

In this example, node 19 is monitoring its neighbors for faulty behaviours using the method proposed in Section IV. Thus the network model knowledge needed is its $2-$hop neighborhood, which consists of 26 states, as opposed to the 236 states of the global network. Using this smaller model, a bank of UIO's was generated according to Section III-A and Section IV.

In the simulations, node 15 exhibits a faulty behaviour after $t = 20\,s$, which is successfully detected by node 19 as seen in Fig. 5. Furthermore, all the residuals corresponding to other neighboring nodes become large while the one for node 15 remains at zero. Following Algorithm 1, node 15 is then detected and identified as the faulty node.

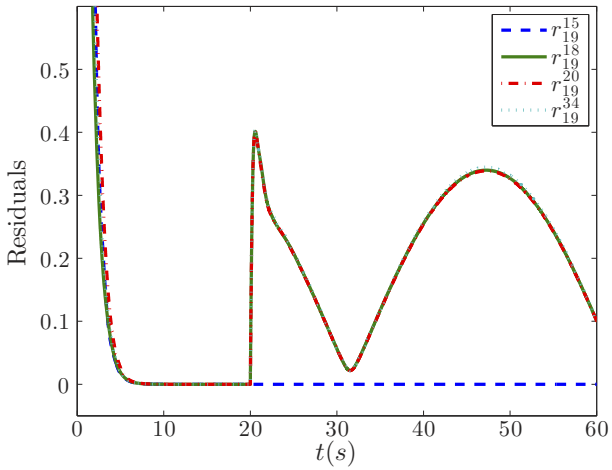


Fig. 5. Residuals generated by the UIO bank at node 19 to detect node faults in $\mathcal{N}_{19}$. A sinusoidal fault is injected by node 15 after $t = 20\,s$. The fault in node 15 is successfully detected and isolated.

B. Faulty edge detection

Here we consider the case where node 15 monitors all its edges as proposed in Section III-B. Note that in power networks the edges represent physical couplings and thus edge faults correspond to parameter faults described in Definition 3.2. We consider the scenario where the system is at equilibrium when the transmission line between nodes 15 and 33 is removed at $t = 5\,s$, which is modeled as $f^w_{15,33}(t) = -w_{15,33}$. This perturbation drives the system to another equilibrium point, enabling us to monitor the state trajectories and locate the faulty edge.

The residuals generated by the observers at node 15 are presented in Fig. 6. As one can see, all the residuals diverge from zero except the one corresponding to the edge between nodes 15 and 33, hence the fault is successfully detected and isolated.

VII. Conclusions and Future Work

The distributed FDI scheme proposed in [22] was extended for detecting and isolating faults in edges of a network. Additionally, the distributed FDI scheme designed using a given initial network model was shown to be robust to the addition or removal of edges. Namely, fault detection can be achieved using this scheme by choosing suitable thresholds, provided that the proximity graph of the monitoring nodes remains constant. Later we establish the minimum measurements required to be able to not only detect but also isolate the faulty nodes by each agent where the only model information they have is a local network model. Then a solution to reduce the computational complexity of the distributed FDI scheme was

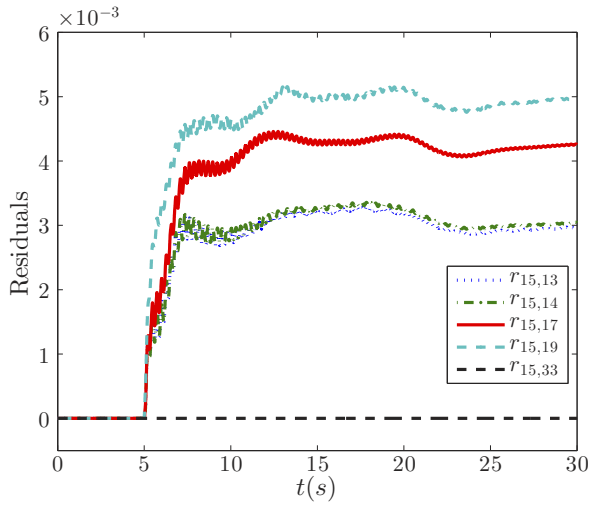


Fig. 6. Residuals generated by the UIO bank at node 15 to detect edge faults. The edge between nodes 15 and 33 is removed at $t = 5\,s$. The edge fault is successfully detected and isolated.

proposed, where the solution lowers the number of monitoring nodes. Numerical result demonstrating the effectiveness of the proposed solutions were presented, taking the IEEE 118 bus power network as an example. As motivated by the example, the proposed methods can be fused to design a scalable and resilient distributed FDI architecture that achieves local fault detection and isolation despite unknown perturbations outside the local subsystem.

Future work includes the design and analysis of the proposed FDI scheme under practical scenarios. In particular, the observer design must account for noise in the system dynamics and measurements. Moreover, it should be devised to ensure good performance of the FDI scheme with respect to relevant metrics such as the detection delay, false alarm rate, and probability of misdetection. In addition to the observer design, the particular choice of thresholds greatly impacts the resulting performance and should also be addressed in future work.

REFERENCES

[1] A. A. Cárdenas, S. Amin, and S. S. Sastry, "Secure control: Towards survivable cyber-physical systems," in *First International Workshop on Cyber-Physical Systems (WCPS2008)*, Beijing, China, June 2008, pp. 495–500.

[2] E. Bompard, C. Gao, R. Napoli, A. Russo, M. Masera, and A. Stefanini, "Risk assessment of malicious attacks against power systems," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 39, no. 5, pp. 1074–1085, Sept. 2009.

[3] C. Rieger, D. Gertman, and M. McQueen, "Resilient control systems: Next generation design research," in *Human System Interactions, 2009. HSI '09. 2nd Conference on*, May, pp. 632–636.

[4] (2004, Apr.) Final report on the August 14th blackout in the United States and Canada. U.S.-Canada Power System Outage Task Force. Accessed Aug. 27, 2014. [Online]. Available: http://energy.gov/sites/prod/files/oeprod/DocumentsandMedia/BlackoutFinal-Web.pdf

[5] (2009, Nov.) Cyber war: Sabotaging the system. CBSNews. Accessed Aug. 27, 2014. [Online]. Available: http://www.cbsnews.com/news/cyber-war-sabotaging-the-system-06-11-2009/

[6] S. Gorman. (2009, Apr.) Electricity grid in U.S. penetrated by spies. The Wall Street Journal. Accessed Aug. 27, 2014. [Online]. Available: http://online.wsj.com/news/articles/SB123914805204099085

[7] N. Lynch, *Distributed Algorithms*, 1st ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1997.

[8] S. Amin, A. Cárdenas, and S. Sastry, "Safe and secure networked control systems under denial-of-service attacks." in *Hybrid Systems: Computation and Control*. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, April 2009, pp. 31–45.

[9] Z. G. Hou, L. Cheng, and M. Tan, "Decentralized robust adaptive control for the multiagent system consensus problem using neural networks," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 39, no. 3, pp. 636–647, June 2009.

[10] S. Sundaram and C. N. Hadjicostis, "Distributed function calculation via linear iterative strategies in the presence of malicious agents," *IEEE Transactions on Automatic Control*, vol. 56, no. 7, pp. 1495–1508, 2011.

[11] F. Pasqualetti, A. Bicchi, and F. Bullo, "Distributed intrusion detection for secure consensus computations," in *Proc. of the 46th IEEE Conference on Decision and Control*, New Orleans, LA, USA, Dec. 2007, pp. 5594–5599.

[12] C. Rieger, "Notional examples and benchmark aspects of a resilient control system," in *3rd International Symposium on Resilient Control Systems (ISRCS)*, Aug. 2010, pp. 64–71.

[13] M. A. Massoumnia and G. C. Verghese, "Failure detection and identification," *IEEE Transactions on Automatic Control*, vol. 34, pp. 316–321, 1989.

[14] J. Chen and R. J. Patton, *Robust Model-Based Fault Diagnosis for Dynamic Systems*. Kluwer Academic Publishers, 1999.

[15] R. Isermann, "Model-based fault detection and diagnosis: status and applications," in *Proceedings of the 16th IFAC Symposium on Automatic Control in Aerospace*, St. Petersburg, Russia, June 2004, pp. 71–85.

[16] S. X. Ding, *Model-based Fault Diagnosis Techniques: Design Schemes*. Springer Verlag, 2008.

[17] I. Hwang, S. Kim, Y. Kim, and C. Seah, "A survey of fault detection, isolation, and reconfiguration methods," *IEEE Transactions on Control Systems Technology*, vol. 18, no. 3, pp. 636–653, 2010.

[18] Z. Han, W. Li, and S. L. Shah, "Fault detection and isolation in the presence of process uncertainties," *Control engineering practice*, vol. 13, no. 5, pp. 587–599, 2005.

[19] E. Scholtz and B. Lesieutre, "Graphical observer design suitable for large-scale DAE power systems," in *Proc. of the 47th IEEE Conference on Decision and Control*, Cancun, Mexico, Dec. 2008, pp. 2955–2960.

[20] M. Aldeen and F. Crusca, "Observer-based fault detection and identification scheme for power systems," in *IEE Proceedings - Generation, Transmission and Distribution*, vol. 153, no. 1, Jan. 2006, pp. 71–79.

[21] F. Pasqualetti, A. Bicchi, and F. Bullo, "Consensus computation in unreliable networks: A system theoretic approach," *IEEE Transactions on Automatic Control*, vol. 57, pp. 90–104, 2012.

[22] I. Shames, A. Teixeira, H. Sandberg, and K. Johansson, "Distributed fault detection for interconnected second-order systems," *Automatica*, vol. 47, no. 12, pp. 2757 – 2764, 2011.

[23] R. Ferrari, T. Parisini, and M. Polycarpou, "Distributed fault diagnosis with overlapping decompositions: An adaptive approximation approach," *IEEE Transactions on Automatic Control*, vol. 54, pp. 794–799, 2009.

[24] Q. Zhang and X. Zhang, "Distributed fault detection and isolation for multimachine power systems," in *2012 IEEE/ASME Int. Conf. on Mechatronics and Embedded Systems and Applications (MESA)*, July 2012, pp. 241–246.

[25] A. Teixeira, H. Sandberg, and K. H. Johansson, "Networked control systems under cyber attacks with applications to power networks," in *Proc. of the American Control Conference (ACC)*, Baltimore, MD, USA, July 2010, pp. 3690–3696.

[26] W. Ren and E. Atkins, "Distributed multi-vehicle coordinated control via local information exchange," *Int. J. Robust Nonlinear Control*, vol. 17, pp. 1002–1033, 2007.

[27] J. Qin, W. X. Zheng, and H. Gao, "Coordination of multiple agents with double-integrator dynamics under generalized interaction topologies," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 42, no. 1, pp. 44–57, Feb. 2012.

[28] P. Kundur, *Power System Stability and Control*. McGraw-Hill Professional, 1994.

[29] P. Barooah and J. Hespanha, "Graph effective resistance and distributed control: Spectral properties and applications," in *Proc. of the 45th IEEE Conference on Decision and Control*, San Diego, CA, USA, Dec. 2006, pp. 3479–3485.

[30] F. Grandoni, "A note on the complexity of minimum dominating set," *J. Discrete Algorithms*, vol. 4, no. 2, pp. 209–214, July 2006.

[31] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER's extensible optimal power flow architecture," in *Proc. IEEE Power and Energy Society General Meeting*, July 2009, pp. 1–7.