



This is an electronic reprint of the original article. This reprint may differ from the original in pagination and typographic detail.

Doostmohammadian, M.; Zarrabi, H.; Charalambous, T.

Sensor fault detection and isolation via networked estimation: rank-deficient dynamical systems

Published in: International Journal of Control

DOI: 10.1080/00207179.2022.2117084

Published: 01/01/2023

Document Version Publisher's PDF, also known as Version of record

Published under the following license: CC BY

Please cite the original version:

Doostmohammadian, M., Zarrabi, H., & Charalambous, T. (2023). Sensor fault detection and isolation via networked estimation: rank-deficient dynamical systems. *International Journal of Control*, *96*(11), 2853-2870. https://doi.org/10.1080/00207179.2022.2117084

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.





ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tcon20

Sensor fault detection and isolation via networked estimation: rank-deficient dynamical systems

M. Doostmohammadian, H. Zarrabi & T. Charalambous

To cite this article: M. Doostmohammadian, H. Zarrabi & T. Charalambous (2023) Sensor fault detection and isolation via networked estimation: rank-deficient dynamical systems, International Journal of Control, 96:11, 2853-2870, DOI: <u>10.1080/00207179.2022.2117084</u>

To link to this article: https://doi.org/10.1080/00207179.2022.2117084

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



6

Published online: 30 Aug 2022.

(
1	

Submit your article to this journal \square

Article views: 610



View related articles 🗹

🕨 View Crossmark data 🗹

Taylor & Francis

OPEN ACCESS

Sensor fault detection and isolation via networked estimation: rank-deficient dynamical systems

M. Doostmohammadian (Da,b, H. Zarrabi^c and T. Charalambous^b

^aFaculty of Mechanical Engineering, Semnan University, Semnan, Iran; ^bSchool of Electrical Engineering, Aalto University, Espoo, Finland; ^cIran Telecommunication Research Center (ITRC), Tehran, Iran

ABSTRACT

This paper considers model-based fault detection of large-scale (possibly rank-deficient) dynamic systems. Assuming only global (and not local) observability over a sensor network, we introduce a single time-scale networked estimator/observer. Sensors take local outputs/measurements of system states with partial observability and share their information (including estimation and/or output) over a communication network, and gain distributed observability. We define the conditions on the network structure ensuring distributed observability and stabilising the error dynamics. However, system outputs are prone to faults and uncertainties, which affect the state estimation of all sensors as a consequence of communicating (possibly) faulty data. From the cyber-physical-systems (CPS) perspective, such faults add bias to the data transferred from the physical layer (dynamic system) to the cyber layer (sensor network). In this work, we propose a localised fault detection and isolation (FDI) mechanism at sensors to secure distributed estimation. This protocol enables every sensor to locally identify the possible fault at the sensor measurement, and, via local detection and isolation, to prevent the spread of biased/faulty information over the network. This distributed isolation and localisation of fault follows from our partial observability assumption instead of full observability at every sensor. Then, other sensors can estimate/track the system by using observationally-equivalent output information to recover for possible loss of observability. In particular, we study rank-deficient systems as they are known to demand more information-sharing, and thus, are more vulnerable to the spread of possible faults over the network. One challenge is the detection of faults in the presence of system/output noise without making (simplifying and unrealistic) upper-bound assumptions on the noise support. We resolve this by adopting probabilistic threshold designs on the residuals. Further, we show that additive faults at rank-deficiency-related outputs affect the residuals at all sensors, a consequence that mandates more constraints on the (distributed) FDI strategy. We address this problem by constrained LMI design of the feedback gain matrix. Finally, we design q-redundant distributed estimators, resilient to isolation/removal of up to q number of faulty sensors, and further, we consider thresholding residual history over a sliding time-window, known as the stateful FDI.

1. Introduction

Designing sensor networks resilient to faults or attacks, has motivated simultaneous estimation and FDI via Linear Dynamical State-space (LDS) models (Barboni et al., 2020; Giraldo et al., 2018; Yılmaz et al., 2016). This is to detect possible anomalies in the system outputs and to avoid biased (and unstable) tracking of the system states. In the CPS setup, such faults appear at the connection from the physical layer to the cyber layer, where well-designed detection strategies can significantly improve CPS reliability. This paper provides a "distributed" (or networked) estimation and FDI protocol to locally detect possible output faults with no need for central coordination. In particular, secure networked estimation with application to CPS (Doostmohammadian et al., 2020; Nweke et al., 2021) and Internet-of-Things (IoT) (Chen et al., 2018) is of interest in this paper. This is motivated by many existing large-scale applications, such as smart-grid monitoring (Abbaszadeh, 2019; Khan & Doostmohammadian, 2011; Khan & Stanković, 2013) and social network inference (Doostmohammadian et al., 2021).

In networked/distributed estimation, a collaborating group of (possibly spatially distributed) sensors tracks the state of the (large-scale) dynamical system with partial observations, local data-processing, and information-sharing over a communication network. The main concern in networked estimation is that the sensors' observations are prone to faults, anomalies, or can be even manipulated by adversaries/malicious-threats. Sharing such faulty data over the sensor network affects the estimation performance of some (or all) other sensors. Our remedy is to design a distributed and local FDI strategy to detect faulty sensors and, further, to reconfigure the sensor network setup by removing the faulty sensors and recovering the loss of observability. In contrast to the existing centralised solutions (e.g. joint estimation and detection in Yılmaz et al., 2016), various networked estimation scenarios are considered in the literature, ranging from multi time-scale (MTS) methods (Battilotti & Cacace, 2021; He, Hu, et al., 2019; He, Ren, et al., 2019; He et al., 2021; Olfati-Saber, 2009) to single time-scale (STS) approaches (Azizi & Khorasani, 2014; Battis-

CONTACT M. Doostmohammadian 🔯 doost@semnan.ac.ir, mohammadreza.doostmohammadian@aalto.fi

This article has been republished with minor changes. These changes do not impact the academic content of the article.

 $\ensuremath{\mathbb{C}}$ 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

ARTICLE HISTORY

Received 21 October 2020 Accepted 19 August 2022

KEYWORDS

Distributed observability; fault detection and isolation; networked estimation; graph contraction; system digraph

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

telli et al., 2012; Boukhobza et al., 2009; Kar et al., 2012; Lopes & Sayed, 2008; Park & Martins, 2012; Ren & Al-Saggaf, 2017). More recently, the protocols are further developed to detect possible attacks/faults in a distributed way (Deghat et al., 2019; Guan & Ge, 2017; He, Ren, et al., 2019; He et al., 2021; Mitra & Sundaram, 2018). Such distributed FDI methods are a significant improvement over the classic centralised FDI counterparts (Chong et al., 2015; Davoodi et al., 2014; Giraldo et al., 2018; Kodakkadan et al., 2017; Lee et al., 2015; Li & Jaimoukha, 2009; Li et al., 2012; Pajic et al., 2015; Rahimian & Preciado, 2015; Rank & Niemann, 1999; Zhang et al., 2021) since the fault can be detected and isolated locally with no help of centralised supervision. Such strategies prevent single-node-of-failure and, further, can be improved by changing the network connectivity; for example, Doostmohammadian et al. (2018) considers costoptimal network design while addressing distributed observability. A practical approach is to add observability redundancy (Lee et al., 2019; Mitra et al., 2021) to make the estimation resilient to the isolation of a few faulty sensors (or sensor failures). In this direction, some works based on structured systems theory are considered in centralised cases, e.g. the centralised graph-theoretic FDI on system digraphs (Commault et al., 2006) and preventive structural design to avoid the so-called zero-dynamics attacks (Weerakkody et al., 2017). Structural-based observability analysis of distributed CPS with applications in distributed estimation and networked control system is discussed in Doostmohammadian and Khan (2020). In some literature, Static Linear State-space (SLS) solutions are considered with no knowledge of the system dynamics. Such centralised SLS scenarios are mostly data-driven using machine learning and binary classification tools (e.g. supportvector-machines) to classify faulty output data from normal data (Abbaszadeh, 2019; Kurt et al., 2018). In general, as compared to LDS solutions, SLS solutions may need more system outputs (in theory as many as system states) to gain system observability. This is also referred to as *static observability* as compared to dynamic observability, where the difference is more significant in large-scale applications with an increasing number of states. Overall, distributed LDS detection scenarios are more compelling in large-scale as they do not require a large number of outputs/sensing-resources and long-range costly (even infeasible) coordination with the centralised detection unit. Further, SLS models only aim for FDI, while LDS models can aim for joint FDI and estimation by knowledge of the system dynamics with possible applications in, e.g.satellite attitude control system (Nasrolahi & Abdollahi, 2018) and Cyber-Physical-Energy-Systems (CPES) (Ilić et al., 2010) as shown in Figure 1.

Main contributions: We propose a joint estimation, detection, and isolation method in a 'distributed setup'. A group of sensors tracks the entire state of a large-scale dynamical system with system matrix *A*, while the system is partially observable to any sensor. Sensors share their estimates and/or measurements to 'gain observability over the network' (referred to as *distributed observability*). To overcome this challenge, some works propose a multi time-scale (MTS) distributed estimation (Battilotti & Cacace, 2021; He, Hu, et al., 2019; He, Ren, et al., 2019; He et al., 2021; Olfati-Saber, 2009), where sensors need to perform *large number* of communications and consensus iterations



Figure 1. This figure shows a possible application of distributed fault detection and state estimation in CPES (Ilić et al., 2010) (in contrast to the centralised solutions). A geographically distributed sensor network monitors a renewable-energy grid including wind and solar-farms. Localized estimation and detection strategies enable distributed data-processing (e.g. in cloud-based architecture) and monitoring with no need for central coordination. This prevents a single-node-of-failure and global shut-down of the entire grid by localising the *detection* and *isolation* of faulty assets.

over the sensor network between every two consecutive timesteps of the system dynamics. Consequently, with more iterations than the sensor network diameter, every sensor receives information of every other sensor and, therefore, the underlying system becomes observable to all sensors between every two system time-steps. This solution imposes high communication load/traffic over the sensor network and fast costly processing/communication units. In particular, this is more challenging for large-scale CPS as the number of sensors and either the sensor network diameter or linking grow larger. To avoid this, single time-scale (STS) distributed estimation protocols are proposed (Azizi & Khorasani, 2014; Boukhobza et al., 2009; Kar et al., 2012), where sensors perform only one step of communication and consensus update at every system time-step (i.e. at the same time-scale). However, these works assume the system is observable in the direct neighbourhood of all (or some) sensors. This requires densely-connected sensor networks where every sensor is directly linked to many other sensors. Both mentioned solutions are costly in terms of either the network traffic or the need for fast communication and computation units, or both. On a large-scale, these might be even infeasible as they require (i) very fast long-range communication and data-processing or (ii) high network traffic at every timestep that may increase congestion, latency, and packet losses. In this paper, we provide a single time-scale (STS) protocol with no local observability assumption at any sensor node, with no need for costly fast communication/processing facilities, and less communication requirement over the sensor network (to reduce the mentioned networking issues). The proposed STS protocol is significant in the following aspects: (i) Unlike (Battistelli et al., 2012; Doostmohammadian et al., 2018; Lopes & Sayed, 2008; Park & Martins, 2012) and similar to Deghat et al. (2019), Mitra and Sundaram (2018), it is not assumed that the system is full-rank or self-damped, and the solution is valid for general possibly rank-deficient systems; (ii) Unlike (Azizi & Khorasani, 2014; Boukhobza et al., 2009; Kar et al., 2012) and similar to Deghat et al. (2019) and Mitra and Sundaram (2018),

no local observability in the neighbourhood of any sensor is assumed to reduce network connectivity and communication traffic/load on sensors.

The performance of the proposed STS protocol is analysed under possible sensing faults. We adopt a residual-based FDI scenario to detect and isolate possible faults (or bias) at state measurements by sensors (considering possible system rankdeficiency). Recall that localisation of faults is more challenging here as the system is not necessarily observable in the neighbourhood of any sensor. This follows our partial (and not full) observability assumption in the local neighbourhood of sensors, however, it demands more complex residual analysis and more constraints on the observer gain design. From structured systems theory perspective (as in Commault et al., 2006), the adjacency graph (or system digraph) of such possibly rank-deficient systems contains the so-called *contraction components*¹ (dual of the dilation components for controllability Doostmohammadian, 2019; Liu et al., 2011). Outputs of state nodes in contractions play key roles in the (distributed) structural observability (Doostmohammadian & Khan, 2020). In this direction, this paper advances the current state-of-the-art in distributed FDI by addressing how rank-deficiency affects networked estimation and detection. We provide a sensor classification to divide sensors into three groups based on their state measurements and the system digraph structure. One class of sensors recovers the system rank-deficiency (with outputs of contractions), the second class recovers the output-connectivity (with outputs of parent-SCCs), and the third class is auxiliary or unnecessary for observability. We show that for the first class any measurement/output fault affects the residuals at all sensors (making it harder to isolate the fault), while for the other two classes the fault only affects the residual at the faulty sensor. This demands different detection and isolation methods at different classes. In this direction, this work extends and generalises our previous works (Doostmohammadian et al., 2021; Doostmohammadian & Meskin, 2020) in which only full-rank dynamical systems are considered. In the proposed STS estimation protocol, the system rank-deficiency is addressed by adding a step of measurement sharing (or innovation-update) over the sensor network. In this work, probabilistic thresholds are designed (in contrast to deterministic thresholds) on the residuals of the noise-corrupted sensors to detect possible additive faults. Stateful measures (also known as distance measures) considering the *history* of the residuals over a *sliding time-window* can also be considered as in Doostmohammadian et al. (2021). The faulty sensor needs to be removed (or isolated) to avoid cascading faulty-data over the distributed estimation network. Further, to recover for distributed observability, a graph-theoretic algorithm is proposed to replace the faulty sensor with an observationally-equivalent one. In another viewpoint, using the notion of *q*-redundant observability, we propose a mechanism to design *q-redundant* distributed estimators/observers, where the observability (and observer error stability) is preserved after removal/isolation of up to q number of failed/faulty sensors. The main contributions of this paper are summarised as follows:

• We propose joint (STS) distributed estimation and (LDS) detection over sensor networks, with no assumption on the

system rank and the local observability at any sensor. As mentioned above in detail, a protocol in this framework outperforms existing ones in terms of needed communication traffic, network connectivity, and data-processing load on sensors.

- This work advances the previous works on full-rank system dynamics (Doostmohammadian et al., 2021; Doostmohammadian & Meskin, 2020), to rank-deficient models. We show that system rank-deficiency imposes more outputs and network connectivity on the distributed estimation network. This makes the detection strategy more challenging in terms of extra constraints on the LMI gain design as compared to full-rank systems in Doostmohammadian et al. (2021), Doostmohammadian and Khan (2013a) and Khan and Jadbabaie (2011).
- Our proposed graph-theoretic sensor classification makes the detection/isolation of faults more simple via structural constraints on the LMI gain design. For example, we add different LMI constraints for each class of sensors, which makes this work different from our previous works (Doostmohammadian et al., 2021; Doostmohammadian & Meskin, 2020).
- Unlike most literature on fault/attack detection (Chong et al., 2015; Kar & Moura, 2011; Kodakkadan et al., 2017; Lee et al., 2015; Pajic et al., 2015; Rank & Niemann, 1999), we do not assume that the system and/or measurement noise are of bounded support. We avoid this simplifying assumption by considering Gaussian noise with no upper-bound. This is more realistic as compared to the existing literature, and results in our probabilistic threshold design. We clearly define the probability of false-alarm and false-negative for the proposed probabilistic thresholds.
- We consider an observability recovery method based on the • concept of observational-equivalence (Doostmohammadian & Khan, 2016; Doostmohammadian et al., 2018). In particular, we provide a graph-theoretic algorithm to replace a faulty sensor (of each class) with an observationallyequivalent one of the same class. This is closely tied with the concept of redundant observability, relating a security index for attack detectability/recovery to observabilitypreserving (Lee et al., 2019). In this direction, a graphtheoretic algorithm is proposed in this paper to design q-redundant networked observers/estimators, i.e. distributed observers robust to failure or removal of up to q sensors. In other words, after removal of up to q number of faulty sensors (and cutting their network links), the remaining sensors can still estimate the system over the (reduced-connectivity) sensor network with stable steady-state error. Similarly, resilient solutions to the removal of up to q communication links over the network can be considered.
- We extend the stateful detection methods in Doostmohammadian et al. (2021) to rank-deficient systems. We introduce distance measures based on residual history and probabilistic thresholds via χ^2 CDF to improve the detection probability and reduce the false-alarm rate.

Paper organisation: Section 2 gives the preliminaries and states the problem. Section 3 describes the networked estimation protocol. In Section 4, the fault detection scenario and

sensor classification are presented. Section 5 presents the algorithms for (i) fault compensation scenario to recover for loss of observability and (ii) designing *q*-redundant distributed estimators. Section 6 provides an illustrating example and comparison with recent literature. Finally, Section 7 concludes the paper.

2. The framework

In this work, we consider noise-corrupted linear systems in discrete-time as,

$$\mathbf{x}_{k+1} = A\mathbf{x}_k + \mathbf{v}_k \tag{1}$$

where $\mathbf{x}_k \in \mathbb{R}^n$ represents the system state (vector) at time-step $k, A = [a_{ij}] \in \mathbb{R}^{n \times n}$ is the system matrix, and $\mathbf{v}_k = \mathcal{N}(0, Q)$ is the system noise with covariance $Q \in \mathbb{R}^{n \times n}$. Note that in this paper we make no assumption on the rank of the system matrix A and it could be rank-deficient. Also, to avoid trivial solutions, we make no assumption on the stability of the system, i.e. it is possible that $\rho(A) > 1$, where $\rho(A)$ denotes the spectral radius of matrix A. System outputs are taken by a group of N sensors as,

$$\begin{pmatrix} \mathbf{y}_k^1 \\ \vdots \\ \mathbf{y}_k^N \end{pmatrix} = \begin{pmatrix} C_1 \\ \vdots \\ C_N \end{pmatrix} \begin{pmatrix} \mathbf{x}_k^1 \\ \vdots \\ \mathbf{x}_k^n \end{pmatrix} + \begin{pmatrix} \boldsymbol{\zeta}_k^1 \\ \vdots \\ \boldsymbol{\zeta}_k^N \end{pmatrix} + \begin{pmatrix} \mathbf{f}_k^1 \\ \vdots \\ \mathbf{f}_k^N \end{pmatrix}$$

where C_i is the measurement/output² matrix at sensor *i*. The above can be written in brief as,

$$\mathbf{y}_k = C\mathbf{x}_k + \boldsymbol{\zeta}_k + \mathbf{f}_k, \qquad (2)$$

where $\boldsymbol{\zeta}_k = \mathcal{N}(0, R)$ is the measurement noise which is assumed to be independent at each sensor. Matrix R is the covariance matrix, which is a diagonal matrix of the variances, and \mathbf{f}_k represents additive fault at sensors. We make standard assumptions on Gaussianity and statistically independency of the noise terms as $\mathbb{E}(\boldsymbol{v}_k \boldsymbol{v}_m) = 0$ and $\mathbb{E}(\boldsymbol{\zeta}_k \boldsymbol{\zeta}_m) = 0$ for all timesteps $k \neq m$. Further, without loss of generality, we assume that every sensor *i* takes the output of one system state (i.e. C_i is a row-vector and y_k^i is a scalar, i.e. $\mathbf{y}_k^i \equiv y_k^i$) and, hence, $\mathbf{y}_k \in \mathbb{R}^N$. Similarly, $\boldsymbol{\zeta}_k^i \equiv \boldsymbol{\zeta}_k^i$ and $\mathbf{f}_k^i \equiv f_k^i$.

Remark 2.1: Typically in fault/attack detection literature (Chong et al., 2015; Kar & Moura, 2011; Kodakkadan et al., 2017; Lee et al., 2015; Pajic et al., 2015; Rank & Niemann, 1999) it is assumed that the noise terms are of bounded-support, i.e. $|\zeta_k^i| < \overline{\zeta}$ and $||v_k|| < \overline{v}$, with $\overline{\zeta}$ and \overline{v} as some upper-bounds. In this paper, we make no such assumptions and consider the noise in its most general form.

2.1 Preliminaries on graph theory

We introduce some graph-theoretic concepts needed for structural analysis of system observability. Denote a graph by $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ with $\mathcal{V} = \{1, ..., n\}$ as the set of nodes and \mathcal{E} as the set of links $i \rightarrow j$ (or pair (i, j)) with $i, j \in \mathcal{V}$. A graph is stronglyconnected (SC) if there exists a directed path from every node to every other node. In a non-SC digraph, define a stronglyconnected-component (SCC) as a component/subgraph in which nodes are strongly-connected to every other node in that component. Among the SCCs, define a parent SCC S_i^p , as the SCC with no outgoing link to any other SCC. The number of parent SCCs is denoted by $|S^p|$ (with $|\cdot|$ as the set cardinality). The decomposition of SCCs and finding their partial order are set via Depth-First-Search (DFS) algorithm (Cormen et al., 2009). Next, define a maximum matching \mathcal{M} as the maximal set of links in \mathcal{E} that share no end-node or beginnode. Define a matched node as the end node of any link in \mathcal{M} . Otherwise, the node is unmatched. Denote the set of unmatched nodes by $\delta \mathcal{M}$. One can find the maximum matching and unmatched nodes via Dulmage-Mendelsohn (DM) decomposition (Dulmage & Mendelsohn, 1958; Murota, 2000). In graph theory, a connected graph G is called κ -connected (or κ vertex-connected) if it remains connected after removing κ (or fewer) nodes. From Menger's theorem, the size of the minimum node-cut (or vertex-cut) between two nodes *i* and *j* (the minimum number of vertices whose removal disconnects *i* and *j*) is equal to the maximum number of pairwise disjoint paths (also referred to as q + 1-linking (Weerakkody et al., 2017)) from i to *j*. In particular, we use *q*-connected digraphs (with *q* pairwise disjoint paths) to design q-redundant observer networks in Section 5.2 along with a simulation example in Section 6. We refer interested readers to Hellwig and Volkmann (2008) for sufficient conditions for designing κ -connected digraphs and to Lau et al. (2009), Sadeghi and Fan (2019) and Umsonst (2019) for a similar concept for survivable network design.

2.2 System digraph representation

For linear system given by (1), define the system digraph $\mathcal{G}_A = \{\mathcal{V}_A, \mathcal{E}_A\}$, as a graph associated to the system matrix A, where $\mathcal{V}_A = \{1, \ldots, n\}$ is the set of *state-nodes* and \mathcal{E}_A is the set of links. There is a link from state-node *i* to node *j*, i.e. $(i, j) \in \mathcal{E}_A$ or $i \rightarrow j$, if $A_{ji} \neq 0$. Many generic system properties including system observability/controllability (Dion et al., 2003; Lin, 1974) are tightly related to the structure of the system digraph. Moreover, one can detect input faults (Commault et al., 2008) or determine the generic/structural rank (or S-rank) of the system matrix (Harary, 1962) using the system digraph \mathcal{G}_A . The following theorem is proved in our previous works (Doostmohammadian & Khan, 2016; Doostmohammadian et al., 2020).

Theorem 2.1: Given the system digraph G_A associated with dynamic system (1), the following state-outputs are necessary for system observability:

- (i) output of one state node in every parent SCC S_l^p , $l \in \{1, ..., |S^p|\}$.
- (ii) outputs of all the unmatched state-nodes in δM .

Note that the number of unmatched nodes in the system digraph is tightly related with the S-rank of the system. It can be proved that Doostmohammadian and Khan (2013b) and Doostmohammadian et al. (2018):

$$|\delta \mathcal{M}| = n - \text{S-rank}(A). \tag{3}$$

with $|\cdot|$ denoting the set cardinality. Recall that the S-rank of the system matrix is the number of non-zero entries in distinct rows and distinct columns of the matrix *A* (Harary, 1962).

In \mathcal{G}_A , the S-rank equals the maximum size of a disjoint family of cycles spanning all state nodes (Reinschke, 1988). Note that based on (3), the number of unmatched nodes equals the rank-deficiency of the system (matrix). Therefore, there is no unmatched node in the system digraph associated with a fullrank dynamical system, e.g. a self-damped system. It is known that condition (i) in Theorem 2.1 recovers the output connectivity and condition (ii) recovers the system S-rank. Therefore, for full-rank systems (as in Doostmohammadian et al., 2021; Doostmohammadian & Meskin, 2020) condition (ii) is already satisfied.

2.3 q-redundant observability

Recall that observability refers to the possibility of reconstructing the entire system states \mathbf{x}_k via some state measurements (outputs) \mathbf{y}_k over finite time-interval. Denote by $\mathcal{O}_{A,C}$ the observability Gramian matrix defined as,

$$\mathcal{O}_{A,C} = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{pmatrix} \tag{4}$$

Having rank($\mathcal{O}_{A,C}$) = *n* implies that the system (1) is observable by outputs (2) and the states (at every time *k*) can be obtained by solving the following set of linear algebraic equations on the outputs,

$$\begin{pmatrix} \mathbf{x}_k \\ \mathbf{x}_{k+1} \\ \vdots \\ \mathbf{x}_{k+n} \end{pmatrix} = \mathcal{O}_{A,C}^{-1} \begin{pmatrix} \mathbf{y}_k \\ \mathbf{y}_{k+1} \\ \vdots \\ \mathbf{y}_{k+n} \end{pmatrix}.$$

Recall that if the conditions (i)–(ii) in Theorem 2.1 hold on the structural system representation, observability Gramian $\mathcal{O}_{A,C}$ is full-rank for almost all numerical values of the non-zero entries in the system matrix *A* and output matrix *C* (Dion et al., 2003; Doostmohammadian & Khan, 2020).

Definition 2.1 (Lee et al., 2019): Define C_{Ω} as the reduced-size output matrix by removing the outputs associated with the set Ω , i.e. by setting the rows of *C* indexed by Ω to all-zeros. The pair (*A*, *C*) in (1)–(2) is then said to be *q*-redundant observable if the pair (*A*, C_{Ω}) is observable for any $\Omega \subset \{1, \ldots, N\}$ satisfying $|\Omega| \ge N - q$.

The above definition simply implies that the system is observable by using *any subset* of outputs of size n-q (or larger), or it remains observable by removing *any subset* of outputs of size q (or smaller), i.e. rank($\mathcal{O}_{A,C_{\Omega}}$) = n. We use this notion along with the structural observational-equivalence to design q-redundant networked observers/estimators in Section 5.2.

2.4 Problem statement

Given the system and measurements by (1)-(2), for (both centralised and distributed) estimation, it is necessary that the pair

(A, C) be observable. Thus, assuming the observability conditions in Theorem 2.1, we introduce an STS networked estimation protocol that requires only one step of estimate sharing and innovation-update (possibly via sharing outputs) between every two time-steps k and k + 1 of system dynamics. We provide structural sensor classification to derive the necessary conditions for distributed observability and design the network topology and the gain matrix to stabilise the networked estimation error. In the case of non-zero additive faults at sensors, a residual-based fault detection logic is proposed. Possible bias $f_k^i \neq 0$ (deviating the real output from the predicted one) causes the residual (or distance measure) at sensor *i* to exceed certain probabilistic thresholds, which triggers the fault alarm at that sensor. The schematic of the networked estimator/observer and the local fault detectors is shown in Figure 2. In the case of detecting a faulty sensor *i* and isolating it (to prevent the distribution of faulty data over the sensor network), a compensation method is proposed to use observationally-equivalent sensors to recover for loss of observability. We further propose a technique to design q-redundant networked estimators, such that by isolation (or removal) of up to *q* faulty (or failed) sensors the system remains observable (in the distributed sense) to the rest of the sensors over the network.

3. Single time-scale networked estimation

In this section, we propose a single time-scale estimator. At every time-step *k* every sensor sends (to its direct out-neighbours) and receives (from its direct in-neighbours) one packet of data consisting of estimate and/or output/ measurement. Every sensor performs *one step* of (i) averaging (consensus) on the priori estimates (to improve its predictions) and, (ii) measurement-update (known as *innovation*) to refine the estimate by output information, which gives the *posteriori* estimate. This represents a *distributed predict* and *update* mechanism, resembling an alpha-beta-type filter as in the Kalman filter. The innovation phase further helps to find the *residual* between the predicted output and the true output to detect possible output bias (for model-based FDI). The proposed distributed protocol is as follows,

$$\widehat{\mathbf{x}}_{k|k-1}^{i} = \sum_{j \in \mathcal{N}_{\beta}(i)} W_{ij} A \widehat{\mathbf{x}}_{k-1|k-1}^{j},$$
(5)

$$\widehat{\mathbf{x}}_{k|k}^{i} = \widehat{\mathbf{x}}_{k|k-1}^{i} + K^{i} \sum_{j \in \mathcal{N}_{\alpha}(i)} C_{j}^{\top} \left(y_{k}^{j} - C_{j} \widehat{\mathbf{x}}_{k|k-1}^{i} \right).$$
(6)

where $\widehat{\mathbf{x}}_{k|k}^{i}$ is the estimate of system state \mathbf{x}_{k} given all the measurements up to time-step k. Define $\mathcal{N}_{\beta}(i)$ as the inneighbourhood of sensor *i* over which sensors share their estimates and $\mathcal{N}_{\alpha}(i)$ as the in-neighbourhood of sensor *i* over which sensors share their measurements. In this direction, two networks are considered: (i) network \mathcal{G}_{β} for estimate-sharing, and (ii) network \mathcal{G}_{α} for output-sharing. The combination of the two, makes the communication network or the sensor network. Both \mathcal{G}_{β} and \mathcal{G}_{α} need to be designed such that distributed observability is satisfied (as we explain in the rest of this section). Matrix *W* includes the consensus weights for averaging prior estimates, i.e. W_{ij} is the consensus gain at sensor *i* on the prior estimate of



Figure 2. This figure shows the schematic of the proposed localised detection via a network of local estimators. All sensors are subject to noise and possible fault. The networked estimators at sensors cooperatively track the states of the dynamical system (e.g. an energy grid) locally while having partial observability. A *localized* detector is embedded at every sensor, which triggers the alarm in case the residual is over a probabilistic threshold (with specified false-alarm (false-positive) and false-negative rates).

the in-neighbour sensor *j*. The structure of *W* follows the graph topology (or structure) of \mathcal{G}_{β} , while the entries satisfy rowstochasticity (e.g. see Charalambous & Hadjicostis, 2013; Xiao et al., 2005) to ensure the consensus nature of Equation (5). We make no constraining assumption on the entries W_{ii} and they follow random consensus weights. In case of using non-random entries, e.g.Metropolis-Hastings fusion rule (Xiao et al., 2005)), the observability should be rechecked numerically (see Doostmohammadian et al., 2021 for more details on this). Define matrix U as the 0–1 adjacency matrix of \mathcal{G}_{α} , i.e. $U_{ij} = 1$ if $j \rightarrow i$ and $U_{ij} = 0$ otherwise. Note that, we assume self cycles (or selfloops) at every node, i.e. $U_{ii} = 1$ and $W_{ii} \neq 0$ for all *i*, since all sensors use their own outputs and priori estimates. The matrix K^{i} , representing the feedback gain for stabilising the estimation error at sensor *i*, needs to fulfil specific constraints for FDI design (discussed in Section 4.1).

Remark 3.1: Recall that in the LDS model, knowing the system dynamics *A* (all sensors), significantly reduces the number of necessary sensors to satisfy system-observability (the so-called Gramian rank-condition in Section 2), while in the *SLS* model with no information of system dynamics more sensors (in theory as many as system states) are needed.

Note that the proposed estimator outperforms the MTS networked estimation proposed in Olfati-Saber (2009), He et al. (2021), He, Ren, et al. (2019), He, Hu, et al. (2019) and Battilotti and Cacace (2021) in terms of real-time capabilities and network communication load at sensors. In these works, sensors need to perform *L* steps of averaging (consensus) and *L* steps of communication between every two steps of system dynamics k-1 and k. This requires the communication and processing units at sensors to be *L* times faster than the sampling of the system dynamics. However, the proposed protocol (5)–(6) only needs 1 step of consensus and communication between time-steps k-1 and k of the system dynamics, see Figure 3.

Remark 3.2: In MTS networked estimation, the number of communication iterations *L* is (at least) more than the diameter d_N of the sensor network. This implies that the information of every sensor reaches every other sensor in the network between every two time-steps k-1 and k, and thus, the system becomes



Figure 3. This figure shows the difference of single and multi time-scale estimators (STS versus MTS). In STS (left) every sensor performs *only* one step of communication and consensus update between two consecutive time-steps k and k + 1 of system dynamics (hence, the same time-scale). In MTS scenario (right) sensors do many steps of communication and consensus update between k and k + 1 to gain local observability (hence, multi time-scale).

observable to every sensor at every time-step k, which makes the solution trivial by imposing high communication/consensus load on sensors.

We need to design the structure of W and U matrices (to address distributed observability) and to find the *structured* gain matrix K to stabilise the estimation error in the steady-state. In this direction, define the estimation error at every sensor i at time-step k as $\mathbf{e}_k^i = \mathbf{x}_{k|k} - \widehat{\mathbf{x}}_{k|k}^i$. This error can be formulated based on the system parameters in Equations (1) and (2) and estimation parameters in Equations (5) and (6) as follows,

$$\begin{aligned} \mathbf{e}_{k}^{i} &= \mathbf{x}_{k} - \left(\widehat{\mathbf{x}}_{k|k-1}^{i} + K^{i} \sum_{j \in \mathcal{N}_{\alpha}(i)} C_{j}^{\top} (y_{k}^{j} - C_{j} \widehat{\mathbf{x}}_{k|k-1}^{i})\right) \\ &= \mathbf{x}_{k} - \left(\sum_{j \in \mathcal{N}_{\beta}(i)} W_{ij} A \widehat{\mathbf{x}}_{k-1|k-1}^{j} \\ &+ K^{i} \sum_{j \in \mathcal{N}_{\alpha}(i)} C_{j}^{\top} \left(y_{k}^{j} - C_{j} \sum_{j \in \mathcal{N}_{\beta}(i)} W_{ij} A \widehat{\mathbf{x}}_{k-1|k-1}^{j}\right)\right). \end{aligned}$$

Substituting (1)–(2) in the above,

$$\mathbf{e}_{k}^{i} = A\mathbf{x}_{k-1} + \mathbf{v}_{k-1} - \left(\sum_{j \in \mathcal{N}_{\beta}(i)} W_{ij}A\widehat{\mathbf{x}}_{k-1|k-1}^{j}\right)$$

$$+ K^{i} \sum_{j \in \mathcal{N}_{\alpha}(i)} C_{j}^{\top} \left(C_{j} \mathbf{x}_{k} + \zeta_{k}^{j} + f_{k}^{j} - C_{j} \sum_{j \in \mathcal{N}_{\beta}(i)} W_{ij} A \mathbf{\hat{x}}_{k-1|k-1}^{j} \right) \right)$$

$$= A \mathbf{x}_{k-1} + \mathbf{v}_{k-1} - \sum_{j \in \mathcal{N}_{\beta}(i)} W_{ij} A \mathbf{\hat{x}}_{k-1|k-1}^{j} - K^{i} \left(\sum_{j \in \mathcal{N}_{\alpha}(i)} C_{j}^{\top} C_{j} (A \mathbf{x}_{k-1} + \mathbf{v}_{k-1}) + C_{j}^{\top} \zeta_{k}^{j} + C_{j}^{\top} f_{k}^{j} - C_{j}^{\top} C_{j} \sum_{j \in \mathcal{N}_{\beta}(i)} W_{ij} A \mathbf{\hat{x}}_{k-1|k-1}^{j} \right).$$

Using the row-stochastic property of the W matrix,

$$A\mathbf{x}_{k-1} = \sum_{j \in \mathcal{N}_{\beta}(i)} W_{ij} A\mathbf{x}_{k-1}$$

Substituting this in the error equation,

$$\mathbf{e}_{k}^{i} = \sum_{j \in \mathcal{N}_{\beta}(i)} W_{ij}A\mathbf{x}_{k-1} - \sum_{j \in \mathcal{N}_{\beta}(i)} W_{ij}A\mathbf{\hat{x}}_{k-1|k-1}^{j}$$

$$- K^{i} \sum_{j \in \mathcal{N}_{\alpha}(i)} C_{j}^{\top}C_{j} \left(\sum_{j \in \mathcal{N}_{\beta}(i)} W_{ij}A\mathbf{x}_{k-1}\right)$$

$$- \sum_{j \in \mathcal{N}_{\beta}(i)} W_{ij}A\mathbf{\hat{x}}_{k-1|k-1}^{j}\right) + \mathbf{v}_{k-1}$$

$$- K^{i} \sum_{j \in \mathcal{N}_{\alpha}(i)} C_{j}^{\top}\zeta_{k}^{j} + C_{j}^{\top}f_{k}^{j} + C_{j}^{\top}C_{j}\mathbf{v}_{k-1}$$

$$= \sum_{j \in \mathcal{N}_{\beta}(i)} W_{ij}A\mathbf{e}_{k-1}^{j} - K^{i} \sum_{j \in \mathcal{N}_{\alpha}(i)} C_{j}^{\top}C_{j} \sum_{j \in \mathcal{N}_{\beta}(i)} W_{ij}A\mathbf{e}_{k-1}^{j} + \boldsymbol{\eta}_{k}^{i},$$
(7)

where $\boldsymbol{\eta}_k^i$ collects the noise and fault terms as,

$$\boldsymbol{\eta}_{k}^{i} := \boldsymbol{\nu}_{k-1} - K^{i} \sum_{j \in \mathcal{N}_{\alpha}(i)} \left(C_{j}^{\top} \boldsymbol{\zeta}_{k}^{j} + C_{j}^{\top} f_{k}^{j} + C_{j}^{\top} C_{j} \boldsymbol{\nu}_{k-1} \right).$$
(8)

Let $\mathbf{e}_k = (\mathbf{e}_k^1; \dots; \mathbf{e}_k^N)$ be the collective error vector. Then,

$$\mathbf{e}_{k} = (W \otimes A - KD_{C}(W \otimes A))\mathbf{e}_{k-1} + \boldsymbol{\eta}_{k} = \widehat{A}\mathbf{e}_{k-1} + \boldsymbol{\eta}_{k}, \quad (9)$$

where $K := \text{blockdiag}(K^i)$ is the block-diagonal gain matrix (to be designed to stabilise the error dynamics (9)) and D_C is defined as,

$$D_C := \begin{pmatrix} \sum_{j \in \mathcal{N}_{\alpha}(1)} C_j^{\top} C_j & & \\ & \ddots & \\ & & \sum_{j \in \mathcal{N}_{\alpha}(N)} C_j^{\top} C_j \end{pmatrix}$$

 Table 1. Comparison between different distributed estimation/detection methods in terms of 'network-connectivity' × 'communication-rate'.

time-scale	STS	STS	MTS
system dynamics communication	$\frac{\text{SLS}}{N(N-1)\times 1}$	$LDS (n_{\alpha}(N-1)+N) \times 1$	$\frac{\text{SLS/LDS}}{2N \times L \text{ with } L \gg 1}$
knowledge of dynamics A	No	Yes	No/Yes

Using these definitions, the collective noise vector η_k is,

$$\boldsymbol{\eta}_k := \mathbf{1}_N \otimes \boldsymbol{\nu}_{k-1} - K D_C (\mathbf{1}_N \otimes \boldsymbol{\nu}_{k-1}) - K \overline{D}_C \boldsymbol{\zeta}_k - K \overline{D}_C \mathbf{f}_k,$$

where $\mathbf{1}_N$ is the column vector of 1s of size N and $\overline{D}_C \triangleq (U \otimes \mathbf{1}_n) \circ (\mathbf{1}_N \otimes C^{\top})$ with 'o' as the entry-wise (or Hadamard) product. Based on Kalman (1960), the error dynamics (9) is stabilizable (in fault-free case) if the pair ($W \otimes A, D_C$) is observable. This condition is referred to as *distributed (or networked) observability* (Doostmohammadian & Khan, 2016). In this direction, the structure of matrices W and U (i.e. the topology of \mathcal{G}_β and \mathcal{G}_α networks) need to satisfy ($W \otimes A, D_C$)-observability. This is discussed in the following lemma.

Lemma 3.1: $(W \otimes A, D_C)$ is structurally observable (referred to as distributed observability) if (i) W is irreducible, i.e. \mathcal{G}_{β} is SC (This results in distributed observability for full-rank systems with rank(A) = n) and (ii), for systems with rank(A) < n, \mathcal{G}_{α} further represents a network of hubs, where sensors with outputs of the unmatched states (referred to as α -sensors in Section 4.1) are the hubs of the network. In other words, every α -sensor j directly shares its output of an unmatched state to every other sensor i, i.e. $j \in \mathcal{N}_{\alpha}(i)$.

Proof: The proof follows from (minimal) sufficient conditions for observability of Kronecker product networks in Doostmohammadian and Khan (2020). Note that the digraph associated with $W \otimes A$ (as the overall distributed system matrix) can be represented as the Kronecker network product (also called the *Tensor graph product*) of the graphs \mathcal{G}_A and \mathcal{G}_β , and then, the lemma directly follows from the results in Doostmohammadian and Khan (2020).

We compared the communication rate and networkconnectivity for different distributed estimation/detection methods in Table 1, considering a large-scale system *A*, where rank(*A*) = $n - n_{\alpha}$ with $n_{\alpha} \ll n$ as the system rank-deficiency.

From Lemma 3.1, for the proposed STS estimator (and the LDS model), the minimum number of links to ensure distributed observability is $n_{\alpha}(n-1)$ on \mathcal{G}_{α} and n on \mathcal{G}_{β} with communication rate of 1 per step ($\mathcal{O}(n_{\alpha}n)$ links). This is illustrated in Section 6 with an example. Note that for large n the STS estimator for the SLS dynamics needs an all-to-all network with $\mathcal{O}(n^2)$ communication links, since the system dynamics A is not (needed to be) known. Note that, later in Section 5, we show that α -sensor i needs to share its output information with sensors that are not observationally-equivalent. Satisfying the above sufficient conditions for distributed observability, one can design the block-diagonal gain matrix K such that matrix \widehat{A} in (9) is a Schur stable, i.e. $\rho(\widehat{A}) < 1$. Further, this

gain matrix K is constrained to be block-diagonal, to address the distributed nature of the networked estimator. We design this K by using iterative cone-complementary optimisation methods (see Ghaoui et al., 1997) as the solution to the following Linear-Matrix-Inequality (LMI),

min **trace**(XY)
s.t.
$$X, Y \succ 0,$$
 K is block-diagonal. (10)
 $\begin{pmatrix} X & \widehat{A}^{\top} \\ \widehat{A} & Y \end{pmatrix} \succ 0,$ $\begin{pmatrix} X & I \\ I & Y \end{pmatrix} \succ 0,$

with stopping criteria $\rho(\widehat{A}) < 1$ or **trace** $(Y_tX + X_tY) < 2nN + \epsilon$ (with predefined ϵ). The iterative algorithm to solve the above LMI can be implemented at the same time-scale of the system dynamics, where the agents update the gain matrix K_k and shared it at each time-step k to address implementations in real-time. For static sensor network and fixed K, however, it is typical to find this K matrix once offline and provide K^i s to sensors before the estimation process. See more information in Doostmohammadian and Meskin (2020), Doostmohammadian and Khan (2013a) and Khan and Jadbabaie (2011).

4. Sensor fault detection: threshold design and sensor classification

This section presents our sensor fault detection logic based on the residual definition and design of probabilistic thresholds via a specific sensor/output classification. Following the methodology in Sundaram (2012), we first define the residuals for all sensors. Given the estimation $\hat{\mathbf{x}}_k^i$ at sensor *i*, define the estimated output at this sensor as $\hat{\gamma}_k^i = C_i \hat{\mathbf{x}}_k^i$. Define the residual as the difference between the original output y_k^i and the estimated (or predicted) output \hat{y}_k^i as,

$$\begin{aligned} \overline{r}_k^i &= y_k^i - y_k^i = y_k^i - C_i \widehat{\mathbf{x}}_{k|k}^i = C_i \mathbf{e}_k^i + \zeta_k^i + f_k^i \\ &= C_i \widehat{A}_i \mathbf{e}_{k-1} + C_i \eta_k^i + \zeta_k^i + f_k^i, \end{aligned}$$
(11)

and the absolute value residual is,

$$r_k^i = |\overline{r}_k^i| = |y_k^i - \widehat{y}_k^i| = |C_i \widehat{A}_i \mathbf{e}_{k-1} + C_i \boldsymbol{\eta}_k^i + \boldsymbol{\zeta}_k^i + f_k^i|, \quad (12)$$

where \widehat{A}_i is the *i*th hyper-row (i.e. the block of rows n(i-1) + 1 to *ni* or $\widehat{A}((i-1)*n+1:i*n,:)$ in MATLAB code) of \widehat{A} . The results of previous section implies that in the absence of faults, i.e. $f_k^i = 0$ for all *i*, the estimation error \mathbf{e}_k^i and the residual r_k^i are bounded steady-state stable for all *i*. Then, one can detect possible faults in case of sufficiently large and biased residuals. The schematic of the proposed fault-detection at every estimator/sensor *i* is presented in Figure 4.

From the Schur stability of \widehat{A} , we skip the first term in (11) and the second term is,

$$C_{i}\boldsymbol{\eta}_{k}^{i} + \boldsymbol{\zeta}_{k}^{i} + f_{k}^{i} = C_{i}\boldsymbol{\nu}_{k-1}$$

$$- C_{i}\boldsymbol{K}^{i}\sum_{j\in\mathcal{N}_{\alpha}(i)} \left(C_{j}^{\top}\boldsymbol{\zeta}_{k}^{j} + C_{j}^{\top}f_{k}^{j} + C_{j}^{\top}C_{j}\boldsymbol{\nu}_{k-1}\right)$$

$$+ \boldsymbol{\zeta}_{k}^{i} + f_{k}^{i}.$$
(13)



Figure 4. Following Figure 2, this figure shows the procedure of fault detection at every sensor/estimator. The difference of the true output and the estimated output gives the residual at sensor/estimator *i*, i.e. $r_k^i = |C_i \mathbf{x}_k^i - C_i \widehat{\mathbf{x}}_k^i|$. The detector at estimator *i* compares the residual with the probabilistic threshold T_{π} (designed in Section 4.2) and raises the alarm if $r_k^i \geq T_{\pi}$. The threshold T_{π} can be designed based on specific false-alarm rate.

In case $f_k^j \neq 0$ for (at least) one sensor $j \in \mathcal{N}_{\alpha}(i)$, the term $C_i K^i C_j^{\top} f_k^j$ adds bias to $C_i \eta_k^i$ and the residual at sensor *i*. This implies that, following the \mathcal{G}_{α} network connectivity in Lemma 3.1, the residual at sensors in the out-neighbourhood of the *faulty* sensor *j* might get biased. In case $\mathcal{N}_{\alpha}(j) = \{j\}$, sensor *j* does not share its output over \mathcal{G}_{α} as $j \notin \mathcal{N}_{\alpha}(i)$ for any $i \neq j$ and $f_k^j \neq 0$ only appears at the residual r_k^j and the fault can be easily isolated. This motivates the sensor classification in the next subsection.

4.1 Sensor classification

Following Lemma 3.1, we classify the sensors based on the type of their state measurement and structure of the system digraph. For a given system digraph G_A , three classes of sensors are as follows:

- *i* ∈ α: every sensor *i* with output of an unmatched state node (in δM). This type of sensors recover system rank-deficiency.
- *i* ∈ β: every sensor *i* with output of a state in a parent SCC S₁^p. This type of sensors recover output-connectivity of the system.
- *i* ∈ γ: every sensor *i* with output of any unnecessary state for observability, i.e. every sensor *i* ∉ α ∪ β.

where the network connectivity requirement of every class (from Lemma 3.1) is defined as follows,³

- Type-α: every α-sensor sends its output *via a direct link* to every other sensor over G_α network.
- Type-β: every β-sensor sends its estimates via a path to every other sensor over network G_β.
- Type-γ: there is no connectivity requirement for a γ-sensor to share its output/prediction. Every γ-sensor only need to receive output information from α-sensors over G_α and estimates from β-sensors over G_β.

In case some contractions and parent SCCs share state nodes, output of that state node is treated as both Type- α and Type- β , see more details in Doostmohammadian (2019) for the dual case of input classification for controllability. Next, assume non-zero additive fault on at least one sensor, e.g. $f_k^j \neq 0$. Recall from Equation (13) that this may affect the residual of only sensor *j* or many other sensors as discussed in the following,

- Type- α : $C_j K^j C_j^\top f_k^j \neq 0$ at the faulty sensor *j* and $C_i K^i C_j^\top f_k^j \neq 0$ at all its out-neighbours $j \in \mathcal{N}_{\alpha}(i)$.
- Type- β and Type- γ : $C_j K^j C_j^\top f_k^j \neq 0$ only at the faulty β/γ -sensor *j* and $C_i K^i C_j^\top f_k^j = 0$ for $i \neq j$, since $\mathcal{N}_{\alpha}(j) = \{j\}$.

In other words, additive fault at every α -sensor (i.e. $\forall i \in \alpha$) affects the residuals at all other out-neighbouring sensors, while fault at any β/γ -sensor only affects the residual at the same sensor. This results in the following fault-detection logic: If only the residual at sensor i is biased (over certain threshold), then sensor i is faulty. If the residual at all sensors is biased, one of the α -sensors is faulty. Since, in the latter case, $f_k^j \neq 0$ at α -sensor j results in $C_i K^i C_j^\top f_k^j \neq 0$ at all out-neighbouring sensors $i, j \in \mathcal{N}_\alpha(i)$, we need to add a new constraint on the gain K in (10). We redesign the matrix K such that $|1 - C_j K^j C_j^\top|$ is sufficiently large for every α -sensor j, while $|C_i K^i C_j^\top|$ is sufficiently small $\forall i \neq j$. This implies that the residual at faulty α -sensor $\forall i \neq j$, and one can detect the faulty α -sensor with greater residual as compared to all other sensors.

Remark 4.1: The design of the gain matrix *K* for distributed FDI on rank-deficient systems is more constrained than full-rank systems. In order to isolate the residuals at every α -sensor *j*, the constrained LMI design in (10) is revised as follows to suppress the cascaded bias at residual of sensors *i* (caused by bias at sensor $j \in \mathcal{N}_{\alpha}(i)$) by constant $0 < \tau < 1$,

min trace(XY)
s.t.
$$X, Y \succ 0,$$
 K is block-diagonal.

$$\begin{pmatrix} X & \widehat{A}^{\top} \\ \widehat{A} & Y \end{pmatrix} \succ 0,$$

$$\begin{pmatrix} X & I \\ I & Y \end{pmatrix} \succ 0,$$

$$\frac{|C_i K^i C_j^{\top}|}{|1 - C_j K^j C_j^{\top}|} < \tau,$$
 $\forall j \in \mathcal{N}_{\alpha}(i), j \neq i$
(14)

with the same stopping criteria as (10).

Note that smaller constant τ implies smaller residual values at neighbouring sensors *i* and, thus, better detection and isolation of faulty α -sensor *j*, while this may affect the convergencetime of the LMI optimisation problem as it reduces the set of possible solutions for *K*. In the next subsection, we design the fault-detection thresholds on the residuals.

4.2 Threshold design

In this subsection, using the noise statistics, we design the thresholds to alarm faults whenever the residual is biased over the threshold. Unlike many literature (Chong et al., 2015; Kar & Moura, 2011; Kar et al., 2011; Kodakkadan et al., 2017; Lee et al., 2015; Pajic et al., 2015; Rank & Niemann, 1999) assuming the noise terms are only supported on a bounded interval as $|\zeta_k^i| \leq \overline{\zeta}$ and/or $\|\boldsymbol{v}_k\| \leq \overline{\nu}$, we do not limit the system/measurement noise to be of bounded support, see Remark 2.1. This is more realistic in real applications of FDI and estimation since the (commonly assumed) Gaussian noise

(as a random variable) is of unbounded support over all real values. This mandates probabilistic threshold design instead of deterministic thresholds in Pajic et al. (2015), Kar et al. (2011), Chong et al. (2015), Lee et al. (2015), Kodakkadan et al. (2017), Kar and Moura (2011) and Rank and Niemann (1999) designed based on the bounds $\overline{\zeta}$ and $\overline{\nu}$.

Let $P_k = \mathbb{E}(\mathbf{e}_k \mathbf{e}_k^{\top})$ and $\Sigma = \mathbb{E}(\boldsymbol{\eta}_k \boldsymbol{\eta}_k^{\top})$. From Equation (9) we have,

$$P_{k} = \widehat{A}P_{k-1}\widehat{A}^{\top} + \Sigma = \widehat{A}^{k}P_{0}(\widehat{A}^{\top})^{k} + \sum_{j=0}^{k-1}\widehat{A}^{j}\Sigma(\widehat{A}^{\top})^{j} \quad (15)$$

Following the Schur stability of \widehat{A} , in the steady-state (as $k \to \infty$),

$$\lim_{k\to\infty}\widehat{A}^k P_0(\widehat{A}^\top)^k\to 0,$$

Therefore,

$$P_{\infty} := \lim_{k \to \infty} P_k = \sum_{j=0}^{\infty} \widehat{A}^j \Sigma (\widehat{A}^{\top})^j$$

Following similar analysis as in Khan and Jadbabaie (2014), one can find a bound on $||P_{\infty}||_2$ as,

$$\|P_{\infty}\|_{2} \leq \sum_{j=0}^{\infty} \|\widehat{A}^{j} \Sigma(\widehat{A}^{\top})^{j}\| \leq \frac{\|\Sigma\|_{2}}{\overline{b}}$$
(16)

for some $\overline{b} < 1$ as a function of $\|\widehat{A}\|_2$. Moreover, in the absence of any faults,

$$\eta_k \eta_k^\top = (I_{Nn} - KD_C) (\mathbf{1}_{NN} \otimes \nu_{k-1} \nu_{k-1}^\top) (I_{Nn} - KD_C)^\top + (K\overline{D}_C) \zeta_k \zeta_k^\top (K\overline{D}_C)^\top$$
(17)

where $\mathbf{1}_{NN}$ denotes the *N* by *N* matrix of all 1s. Applying the \mathbb{E} -operator on both sides of the above equation and recalling that $\mathbb{E}(\mathbf{v}_{k-1}\mathbf{v}_{k-1}^{\top}) = Q$ and $\mathbb{E}(\boldsymbol{\zeta}_k \boldsymbol{\zeta}_k^{\top}) = R$,

$$\Sigma = (I_{Nn} - KD_C)(\mathbf{1}_{NN} \otimes Q)(I_{Nn} - KD_C)^{\top} + (K\overline{D}_C)R(K\overline{D}_C)^{\top}$$
(18)

Taking the 2-norm of the both sides,

$$\|\Sigma\|_{2} \leq \|(I_{Nn} - KD_{C})(\mathbf{1}_{NN} \otimes Q)(I_{Nn} - KD_{C})^{\top}\|_{2} + \|(K\overline{D}_{C})R(K\overline{D}_{C})^{\top}\|_{2} \leq \|I_{Nn} - KD_{C}\|_{2}^{2}N\|Q\|_{2} + \|K\|_{2}^{2}\|\overline{R}\|_{2}$$
(19)

where

$$\overline{R} := \begin{pmatrix} \sum_{j \in \mathcal{N}_{\alpha}(1)} C_{j}^{\top} R_{j} C_{j} & & \\ & \ddots & \\ & & \sum_{j \in \mathcal{N}_{\alpha}(N)} C_{j}^{\top} R_{j} C_{j} \end{pmatrix}$$

Let define $a_1 := \|I_{Nn} - KD_C\|_2^2$ and $a_2 := \|K\|_2^2$. Then, using Equation (16),

$$\frac{\|P_{\infty}\|_{2}}{N} \le \frac{a_{1}N\|Q\|_{2} + a_{2}\|\overline{R}\|_{2}}{N\overline{h}} =: \Phi$$
(20)

 $||P_{\infty}||_2$ defines the 2-norm of the covariance of the collective error \mathbf{e}_k . To find the 2-norm of the covariance of \mathbf{e}_k^i at every sensor *i*, $||P_{\infty}||_2$ is scaled by the number of sensors *N* in Equation (20). This equation defines the bound Φ on the error covariance at every sensor *i*. A similar setup is adopted in Acemoglu et al. (2008) for social learning dynamics. As compared to Acemoglu et al. (2008), in case $||Q||_2 \rightarrow 0$ over time (referred to as diminishing innovations) and $N \rightarrow \infty$, we have $\Phi = 0$ implying perfect estimation for any potentially unstable dynamics in contrast to neutrally-stable systems in Acemoglu et al. (2008).

Following Khan and Jadbabaie (2014), for the fault-free case (and large $k \to \infty$ in steady-state), $\mathbb{E}(\mathbf{e}_k^i) = 0$ and $\mathbb{E}(\mathbf{\bar{r}}_k^i) =$ 0. Then, following the Gaussianity assumption on \mathbf{e}_k^i and r_k^i , one can claim that $r_k^i = |C_i e_k^i + \zeta_k^i| < c\Phi + R$ where $c = ||C_i||_1$ denotes the norm 1 (the sum of absolute value) of the output vector (c is also referred to as the measurement gain). Assuming one state measurement at every sensor i, c is equal to the absolute value of the non-zero entry of C_i . Following the notion of confidence intervals, the probability of $r_k^i < 2c\Phi + 2R$ and $r_k^i < 3c\Phi + 3R$ are (more than) 95% and 99%, respectively. This simply implies that, assuming non-zero fault at any sensor *i*, one can introduce probabilistic thresholds on the residual r_k^i to detect the fault. Let define $T_{68\%} = c\Phi + R$, $T_{95\%} = 2c\Phi + 2R$, $T_{99\%} = 3c\Phi + 3R$ as the probabilistic thresholds on the residuals. If the residual at sensor *i* (or more sensors) is biased over $T_{68\%}$, the detector triggers the alarm and declares possible fault, where the probability of *false alarm* (also known as false positive (Yan et al., 2019)) in this case is less than 32%. This implies high-rate (32%) of incorrectly raising the alarm when there are no true faults on the sensors. To reduce this rate, one can choose higher thresholds on the residuals, e.g. T_{99%} with (less than) 1% probability of false alarm. For a given low false-alarm rate \varkappa , the threshold is defined as $T_{\pi} = \sqrt{2} \text{erf}^{-1}(\pi)(c\Phi + R)$ with $\pi =$ $1 - \varkappa$. For example, to have false-alarm rate less than $\varkappa = 3\%$, the threshold is $T_{97\%} = \sqrt{2} \text{erf}^{-1}(0.97)(c\Phi + R) = 2.21(c\Phi + R)$ R). Similar claims can be stated for stronger fault detection thresholds, e.g. $T_{99\%} = 3c\Phi + 3R$. In general, for a given residual value r_{k}^{i} the highest probability of false-alarm (and the associated threshold) can be defined as follows (see the illustration in Figure 5),

$$\pi = \operatorname{erf}\left(\frac{r_k^i}{\sqrt{2}(c\Phi + R)}\right) \to T_{\pi}, \quad \varkappa = 1 - \pi.$$
 (21)

Similar to false-alarm rate, the probability of false-negative (i.e.detector raises no-alarm while indeed there is non-zero fault at the sensor/output (Yan et al., 2019)) can be defined as illustrated in Figure 5, where the normalised residuals in the presence (red curve) and the absence (blue curve) of the fault follow Normal PDFs; the probability that $r_k^i < m(c\Phi + R)$, i.e. \overline{r}_k^i lies in the confidence interval $(-m(c\Phi - R), m(c\Phi + R))$ of the nonfaulty (zero-mean) PDF, is $\operatorname{erf}(\frac{r_k^i}{\sqrt{2}(c\Phi + R)}) = \operatorname{erf}(m/\sqrt{2}) =: \pi$.



Figure 5. This figure shows an example distribution of non-faulty residual (blue curve) versus faulty residual (red curve) at sensor *i*. The green and red vertical lines represent two example residuals \vec{r}_k^i : the green residual is very close to the expected value (zero) of the fault-free PDF and, thus, most likely is due to system/measurement noise; the red residual is far from the expected value (zero) and, thus, faulty with probability more than 68.3% and less than 95.4% (based on the shown confidence intervals as grey vertical lines). More accurate probability of detection and false alarm for this example can be defined as $\pi = \text{erf}(\frac{1.5}{\sqrt{2}}) = 86.7\%$ and $\varkappa = 1 - \pi = 13.3\%$ (for $r_k^i \ge T_{\pi} = 1.5(c\Phi + R)$), respectively. Considering the absolute value of the residual r_k^i , the blue shaded area equals to $\frac{1-\pi}{2}$, representing half of the probability of false-positive (the other half on the left-side is not shown for simplicity), and the red shaded area equals to $\frac{1-\pi}{2}$, representing (approximate) false-negative probability.

Therefore, in the fault-free case, with probability $1 - \pi$, the *absolute* residual $r_k^i \ge m(c\Phi + R) = T_{\pi}$ and the fault detector triggers the alarm (false positive). With similar line of reasoning for faulty PDF (red curve), with probability $\frac{1-\pi}{2}$ the detector does not trigger the alarm (false-negative)⁴. Note that these thresholds closely depend on the parameters in (20). To find tighter upper-bounds on $\|P_{\infty}\|_2$, one can reduce $\rho(\widehat{A})$, $\|I_{Nn} - KD_C\|_2^2$, and $\|K\|_2^2$ by optimal design of the gain matrix K. This, further, improves the convergence rate of the networked estimation error in the fault-free case (which is a direction of our future research). By embedding these thresholds at every sensor node, sensors can detect possible faults locally at the state output with *no need of any centralised decision making*. Our proposed simultaneous networked estimation and FDI is summarised in Algorithm 1.

4.3 Chi-squared detector via the residual history

One can extend the results of this paper to consider the residual's history over a predefined *sliding time-window* θ , instead of residual at every time-step. This is also referred to as the *stateful* detection (Giraldo et al., 2018) via distance measures. Define, the sum of the normalised squared residuals (at sensor *i*) over the sliding time-window θ as,

$$v_k^i = \sum_{m=k-\theta+1}^k \frac{(r_m^i)^2}{c\Phi + R}.$$
 (22)

It is known that the distribution of this distance measure v_k^i follows the *Chi-square* PDF with degree θ , i.e. $\mathbb{E}[v_k^i] = \theta$ (Doost-mohammadian et al., 2021; Greenwood & Nikulin, 1996; Renganathan et al., 2020). Extending the results in Doostmohammadian et al. (2021) to rank-deficient systems, one can design

Algorithm 1: The proposed algorithm for distributed estimation and FDI.

- 1 **Input:** System matrix *A* and digraph \mathcal{G}_A , system output matrix *C*, false-alarm rate \varkappa or detection probability $\pi = 1 \varkappa$.
- 2 Classify the sensors (α , β , or γ) as in Section 4.1;
- 3 Design the networks \mathcal{G}_{α} and \mathcal{G}_{β} (the sensor network) via Lemma 3.1;
- 4 Design the block-diagonal gain matrix *K* via LMI (14) with small $\tau < 1$;
- 5 **begin** at every time-step *k* every sensor *j*:
- 6 Shares estimate $\widehat{\mathbf{x}}_{k-1|k-1}^{j}$ with sensors $\{i|j \in \mathcal{N}_{\beta}(i)\}$ and measurement y_{k}^{j} with sensors $\{i|j \in \mathcal{N}_{\alpha}(i)\}$;
- 7 Finds the priori estimate $\widehat{\mathbf{x}}_{k|k-1}^{j}$ via Equation (5) and posteriori estimate $\widehat{\mathbf{x}}_{k|k}^{j}$ via Equation (6) using the received neighbouring data;
- 8 Finds the residual r_k^j via Equation (11);
- 9 Finds the threshold $T_{\pi} = \sqrt{2} \text{erf}^{-1}(\pi)(c\Phi + R)$ for given $\pi = 1 \varkappa$;
- 10 if $r_k^j > T_\pi$ then

11

Alarm: sensor *j* is faulty with probability π (or false alarm rate \varkappa);

12 **Output:** estimate
$$\widehat{\mathbf{x}}_{k|k}^{j}$$
, fault decision at sensor *j*;

thresholds using the CDF of χ^2_{θ} -distribution as,

$$T_{\pi} = 2\Gamma^{-1}\left(\pi, \frac{\theta}{2}\right) = 2\Gamma^{-1}\left(1 - \varkappa, \frac{\theta}{2}\right), \qquad (23)$$

with \varkappa as the probability of false-alarm and $\Gamma^{-1}(\cdot, \cdot)$ denoting the *inverse regularised lower incomplete gamma function* (Greenwood & Nikulin, 1996). However, such detection method typically raises the alarm after certain delay depending on θ . This is because, instead of the instantaneous residual at one timestep, the *summation* in (22) over a (possibly) large time-window θ is considered and compared with the thresholds; therefore, the deviation of the residuals over *many* time-steps over θ would raise the alarm. This is more illustrated by simulation in Section 6.

5. Sensor fault compensation, observational-equivalence, and *q*-redundant networked observers

5.1 Observationally-equivalent state nodes

In this subsection, we propose a strategy to replace the faulty sensor (detected via the logic in the previous section) with an observationally-equivalent sensor. The concept of *observational-equivalence*, introduced in Doostmohammadian and Khan (2016) and Doostmohammadian et al. (2018), provides a methodology to recover for loss of observability by replacing the biased/faulty output with an equivalent output in terms of observability. This concept is better defined on the system digraph (see Section 2.2). Recall that form Theorem 2.1

outputs/measurements of (i) one state in every parent SCC and (ii) every unmatched node are necessary for observability. All state nodes in the same parent SCC, S_1^p , are observationallyequivalent in the sense that output of any one fulfils the observability condition in Theorem 2.1. On the other hand, the set of unmatched nodes δM and maximum matching M are not unique in general (Murota, 2000). In this direction, a contraction, C_l , is defined as the set of nodes in \mathcal{G}_A such that $|\mathcal{N}(\mathcal{C}_l)| < |\mathcal{C}_l|$, where $\mathcal{N}(\mathcal{C}_l) = \{j|(i,j) \in \mathcal{E}, i \in \mathcal{C}_l\}$. Let the set of all contractions in \mathcal{G}_A be denoted by \mathcal{C} ; for every choice of maximum matching \mathcal{M} , there exists one unmatched node in every contraction C_l and the number of all contractions equals the number of unmatched nodes in \mathcal{G}_A , i.e. $\mathcal{C} = \{\mathcal{C}_1, \ldots, \mathcal{C}_{|\delta \mathcal{M}|}\}$. It is known that every observation/output of a state node in a contraction C_l recovers the system S-rank by 1 (Doostmohammadian et al., 2018). In other words, denoting the outputs of all nodes in C_l by C_{C_l} , we have,

$$\operatorname{S-rank}\begin{pmatrix}A\\C_{\mathcal{C}_l}\end{pmatrix} = \operatorname{S-rank}(A) + 1.$$
(24)

The above holds for any nonempty subset of C_l , and implies the observational-equivalence of all state nodes in C_l . From Section 2.3, output of states in C_l also recover the rank of \mathcal{O}_{A,C_l} by (at least) 1. We refer interested readers to our previous work (Doostmohammadian et al., 2018) for more information. Define the sets $\mathcal{A}_i \subset \alpha$, $\mathcal{B}_j \subset \beta$ with $i \in \{1, \ldots, |\delta\mathcal{M}|\}$, $j \in \{1, \ldots, |\mathcal{S}^p|\}$ as observationally-equivalent sensors with outputs from states in C_i and \mathcal{S}_j^p , respectively. Using the notion of observational-equivalence, our fault compensation logic is as follows,

- Type-α: replace the faulty sensor a ∈ A_l by another sensor a' ∈ A_l (both a and a' possess outputs from the same contraction C_l).
- Type-β: replace the faulty sensor b ∈ B_l by sensor b' ∈ B_l (both b and b' possess outputs from the same parent SCC S^p_l).
- Type-γ: remove the faulty sensor of this type. This type plays no necessary role in system observability.

where the above compensation logic assumes that $|A_i| \ge 2$ and $|B_i| \ge 2$. This assumption says that there exist (at least) two (observationally-equivalent) state nodes in every parent SCC and contraction. The connectivity of the new substitute α or β sensor follows the same connectivity as the removed faulty one (see connectivity conditions in Section 4.1). For Type- γ , there is no need for substitute sensors, and the network connectivity ity of the remaining sensors needs to be adjusted to fulfil the conditions in Section 4.1.

5.2 Proposed algorithm for q-redundant networked observability

Following the definitions in Section 2, the concept of observational-equivalence is in close relation with q-redundant observability. We are interested to design resilient networked estimators to tolerate isolation/removal of faulty sensors (or **Algorithm 2:** The proposed algorithm to recover for networked observability.

1	Input: System di	graph \mathcal{G}_{I}	A, faulty a	sensor <i>j</i> from	
	Algorithm 1.				

- Remove sensor *j* and cut all its communication links over *G_α* and *G_β*;
- 3 if $j \in \alpha$ then
- 4 Find $\{l|l \in \{1, \ldots, |\delta \mathcal{M}|\}, j \in \mathcal{A}_l\};$
- 5 Replace sensor *j* by (observationally-equivalent) sensor $j' \in A_l$;

6 else

7 | **if** $j \in \beta$ then

- 8 Find $\{l | l \in \{1, ..., |S_l^p|\}, j \in B_l\};$
- 9 Replace sensor *j* by (observationally-equivalent) sensor $j' \in \mathcal{B}_l$;

10 Link sensor j' to (and from) out-neighbours (in-neighbours) of (removed) sensor j;

11 Output: substitute observationally-equivalent sensor j' over the network;

failed sensors). In this direction, Algorithm 3 designs qredundant observable estimators. q number of detected faulty sensors (via the FDI logic in Algorithm 1) can be isolated/removed, while the remaining sensor network preserves distributed observability. Note that q is limited by the minimum size of observationally-equivalent sets (i.e. contractions and parent SCCs), and this min size is defined by \overline{q} in Algorithm 3. Similarly, one can extend the results to design q-edge-connected networks which remain SC after removal of (less than or equal to) q number of links. This is referred to as survivable network design (Lau et al., 2009; Sadeghi & Fan, 2019; Umsonst, 2019) and particularly is related to the connectivity requirement of the Type- β sensors. In other words, designing a *q*-edge-connected sensor network ensures strong-connectivity after removal of (up to) q links (or q lost-connectivity/missing-packets), which guarantees the connectivity requirement sensors over \mathcal{G}_{β} . Similarly, for \mathcal{G}_{α} , one can add more links from observationally-equivalent α sensors.

6. Illustrative example and simulation

6.1 Illustrative simulation

We consider a linear system of 12 states with the digraph shown in Figure 6. The link weights (chosen randomly in [0.02 1.1]) represent the non-zero entries of system matrix *A*; for example, link from state 1 to 2 with weight 0.79 implies that $a_{21} =$ 0.79 (a_{21} is the entry at column 1 and row 2 of *A*). The system is considered in the discrete-time form (1) with noise v = $\mathcal{N}(0, 0.01)$. The associated system digraph includes two parent SCCs $S_1^p = \{1, 2, 3\}, S_2^p = \{8, 9\}$ and the set of unmatched nodes is $\delta \mathcal{M} = \{4, 10\}$ and contractions $C_1 = \{4, 8, 9\}, C_2 =$ $\{10, 11\}$. Note that state nodes 8, 9 are shared between C_1 and S_2^p . Following Theorem 2.1, state 1 from S_1^p and state 8 from S_2^p and unmatched states $\{4, 10\}$ are respectively measured by

Algorithm 3: Algorithm to design q-redundant net-						
wo	worked estimators.					
1 I	1 Input: System digraph \mathcal{G}_A , redundancy q					
2 F	Find contractions $C = \{C_1, \ldots, C_{ \delta M }\};$					
3 F	Find parent SCCs $S^p = \{S_1^p, S_2^p, \dots, S_{ S^p }^p\};$					
4 I	Define $\overline{q} = \min_{i \in \{1, \dots, \delta \mathcal{M} \}, j \in \{1, \dots, \mathcal{S}^p \}} \{ \mathcal{C}_i , \mathcal{S}_j^p \} - 1;$					
5 F	Return: max possible redundancy as \overline{q} ;					
6 i	$\mathbf{f} \ \mathcal{C}_i \cap \mathcal{S}_j^p = \emptyset \ for \ i \in \{1, \dots, \delta \mathcal{M} \}, j \in \{1, \dots, \mathcal{S}^p \}$					
	then					
7	For $1 \le q \le \overline{q}$ define $N = (q+1)(\mathcal{S}^p + \delta\mathcal{M})$ as					
	the min number of necessary state outputs for					
	<i>q</i> -redundant observability;					
8	Assign N sensors to $q + 1$ state-outputs from every					
	$C_i \text{ and } S_j^p, i \in \{1, \dots, \delta \mathcal{M} \}, j \in \{1, \dots, S^p \} \text{ and }$					
	design output matrix C;					
9	Define $\mathcal{A}_i, \mathcal{B}_j, i \in \{1, \dots, \delta \mathcal{M} \}, j \in \{1, \dots, \mathcal{S}^p \}$ as					
	the set of $q + 1$ observationally-equivalent sensors					
	(with outputs from the same parent					
	SCC/contraction);					
10 e	lse					
11	Define $\overline{N} = \sum_{i \in \{1, \dots, \delta \mathcal{M} \}, i \in \{1, \dots, S^p \}} \{\mathcal{C}_i \cap \mathcal{S}_i^p \neq \emptyset\}$					
	as the number of non-disjoint parent SCC and					
	contraction sets;					
12	Reduce min number of necessary outputs N (by at					
	most \overline{N}) via taking outputs from the shared state					
	nodes in every set $\{C_i \cap S_i^p \neq \emptyset\}$, see details in					
	Doostmohammadian (2019) Any sensor l with					
	outputs of states in $\{C_i \cap S_i^p \neq \emptyset\}$ is both Type- α					
	and β , i.e., $l \in A_i, l \in B_j$;					
13 $\forall i \in \{1, \dots, \delta \mathcal{M} \}, \forall m \notin \mathcal{A}_i, \forall l \in \mathcal{A}_i: \text{ add a link } l \to m$						
in \mathcal{G}_{α} , i.e., $m \in \mathcal{N}_{\alpha}(l)$;						
14 Design \mathcal{G}_{β} as a $(q+1)$ -connected network of all sensors						
	(see Section 2.1):					

- 15 Apply the networked estimation protocol (5)–(6) over the networks \mathcal{G}_{α} and \mathcal{G}_{β} ;
- **16 Output:** (structured) matrix *C* and networks $\mathcal{G}_{\alpha}, \mathcal{G}_{\beta}$;

sensors β_1, β_2 (Type- β) and α_1, α_2 (Type- α), with measurement noise $\zeta = \mathcal{N}(0, 0.01)$. These outputs fulfil the necessary conditions for (A, C)-observability (see the observability conditions in Doostmohammadian & Khan, 2020). The networks \mathcal{G}_{α} and \mathcal{G}_{β} are designed as given in Figure 6, satisfying distributed observability conditions in Lemma 3.1. The networked estimation protocol follows Equations (5) and (6), where sensors share their estimates over the SC network \mathcal{G}_{β} (solid arrows) and their measurements over the network \mathcal{G}_{α} (dashed arrows). Link weights in \mathcal{G}_{β} are chosen randomly such that the incoming link weights sum to 1 to satisfy the row-stochasticity of consensus matrix W, and for network \mathcal{G}_{α} the link weights are equal to 1. The output gains, i.e. the non-zero entries of the C matrix, are also equal to 1. System is unstable as $\rho(A) = 1.155 > 1$. In \mathcal{G}_{α} , the α -sensors are the hubs of the network and directly send their outputs to other sensors. We choose $\tau = 0.2$ in the LMI gain design (14), so that the residual at a faulty α -sensor (say α_2) be



Figure 6. This figure shows a setup for networked estimation over an example system digraph (left), with outputs of the blue-coloured state nodes taken by the redcoloured "sensor" nodes. The communication network (right) of the sensors includes network \mathcal{G}_{α} (dashed arrows) over which α -sensors share their outputs to every other sensor and SC network \mathcal{G}_{β} (solid arrows) over which sensors share their estimates. See Section 3 for details.

about $\frac{1}{0.2} = 5$ times greater than the residuals at the other non-faulty sensors (say $\alpha_1, \beta_1, \beta_2$). Therefore, the conditions in (14) are as follows,

$$C_i K^i C_j^{\top} | \le 0.2 |1 - C_j K^j C_j^{\top}|, \ j = \alpha_1, \quad \alpha_2, \ j \ne i$$
 (25)

For the sake of space, only K^{α_2} (feedback gain at sensor α_2) is given below as an example and the rest are skipped.

	(0.95	0	0	0		0	0	0
	0	0	0	0		0	0	0
	0.56	0	0	-0.653	3	0	0	0
	0	0	0	0.95		0	0	0
	0	0	0	0		0	0	0
V^{α_2}	0	0	0	0		0	0	0
K - =	0	0	0	0.1338		0	0	0
	0	0	0	0		0	0	0
	0	0	0	0		0	0	0
	0	0	0	0		0	0	0
	0	0	0	0		0	0	0
	0	0	0	0		0	0	0
	0		0	0	0	0)	`	
	-0.033	4	0	0.1584	0	0		
	0		0	0	0	0		
	0		0	0	0	0		
	0.0294		0	0.2730	0	0		
	-0.029	9	0	0.142	0	0		
	0		0	0	0	0		
	0.95		0	0	0	0		
	0.8238		0	0.0697	0	0		
	0		0	0.95	0	0		
	-0.026	5	0	0.1257	0	0		
	0		0	0	0	0	/	

The constrained LMI (14) is solved by MATLAB CVX, and the resulting *K* gives $\frac{|C_i K^i C_j^\top|}{|1-C_j K^j C_j^\top|} < 1 \times 10^{-6}$ for $j = \alpha_2$ and $i = \alpha_1, \beta_1, \beta_2$, respectively. (similarly for $j = \alpha_1$). This implies that the residual due to (sufficiently large) fault value f_k^j at sensor $j = \alpha_2$ is at least 5 times greater than the residual at other sensors $i \neq j$. This helps to easily detect/isolate the fault at sensor $j = \alpha_2$. We assume additive faults at sensors α_2 and β_1 as $f_{k\geq 25}^{\alpha_2} \mathcal{N}(1.2, 0.5)$ and $f_{k\geq 55}^{\beta_1} = 1.2$. In this example, $\overline{b} = 0.0505$, $a_1 = 1.753, a_2 = 1.6551$, and $\|\overline{R}\|_2 = 0.01$. Putting these values in (20), $\Phi = 0.41$ and the thresholds are,

$$T_{68\%} = 0.42, \quad T_{95\%} = 0.84,$$

 $T_{99.7\%} = 1.26, \quad T_{99.99\%} = 1.68$

The sensor residuals via Equation (12) and mean-squareestimation-errors (MSEEs) under the proposed networked estimation protocol (5)–(6) along with the thresholds are shown in Figure 7. One can claim that sensor α_2 and β_1 are faulty with a probability of 95% (for $k \ge 25$ and $k \ge 55$ respectively) and the probability of false alarm is less than 5%. Next, following Section 4.3, we consider stateful detection via the distance measure (22) with sliding time-window of $\theta = 12$ -steps over the residual history. The thresholds (with the same detection probabilities) can be found via (23) as,

$$T_{68\%} = 13.7, \quad T_{95\%} = 21.1,$$

 $T_{99.7\%} = 29.8, \quad T_{99.99\%} = 39.1$

From Figure 7, the Chi-squared detector raises a higher detection probability of 99.7% (lower false-alarm rate 0.3%), however, with certain delays with respect to the exact fault-occurrence times.

Next, we replace the two faulty sensors with new observationally-equivalent sensors via the results in Section 5 to recover for loss of observability. The unmatched state 10 (with output to sensor α_2) belongs to the contraction $C_2 = \{10, 11\}$. We replace α_2 by non-faulty sensor α'_2 with output of state node 11. Similarly, sensor β_1 with output of state node 1 in the parent SCC $S_1^p = \{1, 2, 3\}$ is replaced by non-faulty sensor β'_1 with output of state node 2. This new setup is shown in Figure 8, where the state



Figure 7. This figure shows the residuals and MSEEs at the setup given in Figure 6 for given faults at sensors β_1 and α_2 . The residuals at faulty sensors β_1 and α_2 are over the threshold $T_{95\%}$, implying possible faults with probability more than 95%. For the same faults, the distance measures are over the threshold $T_{99.7\%}$ and Chi-squared detector reveals a higher detection probability of 99.7%; however, the distance measures go over the thresholds with certain delay (i.e. delay in raising the alarm).



Figure 8. This figure shows the setup to compensate for the loss of observability (after isolating/removing faulty sensors α_2 and β_1 in Figure 6). The set of orange-coloured and green-coloured states respectively belong to a parent SCC and a contraction. The faulty sensors α_2 and β_1 with outputs from states 10 and 1 (in Figure 6) are replaced with new sensors α'_2 and β'_1 respectively with outputs of equivalent states 11 and 2 in the system digraph. The new setup ensures system observability and results in bounded stable MSEE as shown in Figure 9.

nodes in the contraction C_2 and parent SCC S_1^p are coloured. Figure 9 presents the performance of the networked estimation protocol under the compensated setup (after replacing the faulty sensors with their equivalents). As it is clear, the MSEE is bounded steady-state stable at all sensors with no (considerable) residual, implying the distributed observability recovery in the new setup by Figure 8.

Note that, for this example system digraph, $\overline{q} = 1$ from Algorithm 3, as minimum size of the contractions and parent SCCs is equal to 2. Thus, one can design a 1-redundant observer as shown in Figure 10. The design follows from Algorithm 3: we have N = 8 and $\overline{N} = 1$ as $C_1 \cap \beta_1 = \{4, 9\}$, and the minimum number of necessary outputs for this case is 7. Therefore, the sensor with output of state 9 is both Type α and β , shown as $\alpha'_1\beta'_1$ in Figure 10. Every sensor receives two direct links from every set $\mathcal{A}_1 = \{\alpha_1, \alpha'_1\}$ and $\mathcal{A}_2 = \{\alpha_2, \alpha'_2\}$. The undirected network \mathcal{G}_β as an undirected cycle of all sensors is 1-connected (see

 Table 2. Parameter values for the MTS protocol in He et al. (2021) and He, Ren, et al. (2019).

L	20	α	0.5	β	0.5
<i>A</i>	1.28	γ	0.2	Ν	4
S	2	b _w	0.1	b _v	0.1
λ0	1.95	η_0	0.1	ρ_{t_0}	0.1

Section 2.1), i.e. by removing any 1 sensor (or any 1 link) the network \mathcal{G}_{β} remains SC. Therefore, after removal/isolation of any α or β sensor (for example sensor α_1), the remaining sensor network still satisfies Lemma 3.1 and Theorem 2.1 (i.e. distributed observability is preserved). Recall that the performance of this 1-redundant networked observer/estimator after removing one sensor from every observationally-equivalent set $\mathcal{A}_1, \mathcal{A}_2, \mathcal{B}_1, \mathcal{B}_2$ is given, for example, in Figures 6 and 7 and in Figures 8 and 9.

6.2 Comparison with recent literature

In this subsection, we compare our proposed methodology with recent works (He et al., 2021) and He, Ren, et al. (2019) as they make similar assumptions on system stability and distributed observability, while most literature considers stable dynamics and local observability in the direct neighbourhood of sensors. He et al. (2021) and He, Ren, et al. (2019) propose an MTS protocol for distributed filtering and fault detection, which requires L steps of consensus between two consecutive steps of system dynamics. The detection logic in He et al. (2021) and He, Ren, et al. (2019) is deterministic and claims to detect faults of a certain magnitude. We run this MTS protocol over the same setup as in the previous subsection (Figure 6) for comparison, using the same system parameters and outputs. The only difference is that in He et al. (2021) and He, Ren, et al. (2019) the sensor network needs to be *undirected* and we used bidirectional \mathcal{G}_{β} with weights $\overline{W} = \frac{W^{\top} + W}{2}$ instead. Table 2 gives the chosen parameters of the MTS protocol. Adopting these parameters the MSEEs at all 4 sensors are shown in Figure 11. According to He et al. (2021) and He, Ren, et al. (2019), the fault is detected whenever the measurement-update at a sensor is over the threshold A. From the figure, one can see that both faults at sensors α_2 and β_1 are detected. However, the detector triggers alarm falsely at sensor β_2 .

The improvements of our proposed distributed strategy over this MTS solution are as follows:



Figure 9. The MSEEs, residuals, and distance measures at all 4 sensors in the recovered setup in Figure 8 are shown here. The steady-state stability of MSEEs at all sensors implies the networked observability of the proposed recovered estimation setup.



Figure 10. (Left) 2 outputs from states in every parent SCC and contraction are considered in the system digraph, so that the pair (A, C) is 1-redundant observable. (Right) The proposed network topology to design 1-redundant networked observer/estimator is shown (following Section 4.1). The networked estimation can tolerate removing/isolating any one faulty sensor. For example, assume S_{α_1} is isolated by cutting all its communications (colored by red). The black-coloured links show the reduced communication network of G_{β} (solid links) and G_{α} (dashed links). The conditions in Lemma 3.1 and Theorem 2.1 hold on the reduced setup, i.e. the system (digraph) is structurally observable (in distributed sense) to *all other sensors* communicating over the reduced network. This 1-redundant networked observer can tolerate isolation/removal of 1 sensor from every observationally-equivalent set A_1 , A_2 , B_1 , B_2 as shown in Figures 6–7 and 8–9.

First (and most important), the MTS strategy in He et al. (2021) and He, Ren, et al. (2019) requires L (in this example 20) consensus and communication iterations between time-steps k and k+1, where L needs to be (at least) more than the diameter d_N of the sensor-network and, further, needs to be large enough for resilient and faulttolerant filtering. $L > d_N$ guarantees that the information of every sensor reaches every other sensor between k and k+1. Thus, the system becomes locally observable to all sensors at all time-steps k via the many iterations of consensus/communication. In contrast, our proposed networked estimator gains distributed observability at the same timescale k via the established network topology and LMI gain design. The MTS approach in He et al. (2021) and He, Ren, et al. (2019) requires (L-times) faster communication facilities along with (L-times) faster computation units to process the received information at every sensor. In largescale systems, e.g. geographically distributed grid monitoring, with long-range communications and sparse network connectivity (large diameter d_N), the MTS approach is more costly and even infeasible. Our proposed strategy demands low-cost equipment with equivalent (and even better) estimation and FDI performance (compare the MSEE performance in Figures 7 and 11). Following Table 1, here the 'network-connectivity' × 'communication-rate' is 8×20 as compared to $(6 + 4) \times 1$ in Figure 6.

• The strategy in He et al. (2021) and He, Ren, et al. (2019) detects the fault but provides no recovery solution for faulty sensors, i.e. all other sensors keep receiving biased information from the faulty ones. Having few averaging iterations (small *L*) may bias the MSEEs at all sensors (see Figure 11). To overcome this, the number of averaging iterations *L* needs to be increased, which, in turn, demands even faster communications/data-processing. In contrast, using our recovery method via observational-equivalence, one can recover the loss of observability and even add redundant sensors to make the system *q*-redundant observable. For example, compare the MSEE performance of Figure 9 (after observability recovery) and Figure 11.



Figure 11. The MSEEs at 4 sensors are shown using the MTS protocol in He et al. (2021) and He, Ren, et al. (2019) via L = 20 steps of averaging/consensus between every two consecutive steps of system dynamics. The measurement-updates (while tracking the unstable system in Figure 6 via the MTS protocol) are given at all 4 sensors. This figure says that the faults at sensor α_2 and β_1 is above the threshold Λ (detected), while the detector raises false alarm at sensor β_2 .

7. Conclusion

A residual-based distributed FDI method over distributed estimation networks and a probabilistic threshold on the residuals is proposed in this paper to detect possible additive faults. We validated our results on an academic simulation example. The performance measures of the proposed joint estimation and FDI algorithm are summarised as follows: (i) no need for high communication/computation rate and costly networking/processing resources as compared to MTS estimators, (ii) less network traffic/connectivity as compared to locally observable STS estimators, (iii) more accurate detection via probabilistic threshold design with no simplifying upper-bound assumption on the noise support (as compared to the deterministic thresholds), and (iv) estimation recovery via adding observationally-equivalent sensors and designing q-redundant and fault-tolerant distributed estimators.

It is worth mentioning that the proposed methodology in this paper is of polynomial-order complexity, which makes it applicable in large-scale systems. The LMI approach to design the gain matrix is of polynomial-order complexity (Nesterov & Nemirovskii, 1994; Ye, 1993). Moreover, the computational complexity of the DFS algorithm is $O(n^2)$ (Cormen et al., 2009) and the computational complexity of the most efficient algorithm for DM decomposition and finding contractions in the system digraph is $O(n^{2.5})$ (Micali & Vazirani, 1980). As a direction of future research, we are aiming to consider possible time-delays over the communication network using the results in Hadjicostis and Charalambous (2014) and Doostmohammadian et al. (2021).

Notes

- 1. The relevant concepts of contractions and parent SCCs (stronglyconnected-components) and other structural observability notions are defined later in Section 2.1.
- 2. Throughout this paper, the terms output and measurement are used interchangeably.
- Note that in this section, we do not consider observationally-equivalent sensors, e.g, with outputs from the same parent SCC S_l^p. This is discussed later in the next section.
- 4. Note that the given probability for false-negative is the approximate value as we consider the *absolute residual* rⁱ_k. This implies *folded* normal distribution, i.e. to fold-over the probability mass to the right of 0 axis by taking absolute value. Then, the exact probability of false-negative is

$$\overline{\varkappa} = \frac{1 - \operatorname{erf}(m/\sqrt{2})}{2} - \frac{1 - \operatorname{erf}(3m/\sqrt{2})}{2} = \frac{\operatorname{erf}(3m/\sqrt{2}) - \operatorname{erf}(m/\sqrt{2})}{2}.$$
 For $m \ge 2$ one can approximate $\overline{\varkappa}$ by $\frac{1 - \pi}{2}$ as $1 - \operatorname{erf}(3m/\sqrt{2}) \approx 2e - 10.$

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

M. Doostmohammadian D http://orcid.org/0000-0003-0959-6608

References

- Abbaszadeh, M. (2019). System and method for anomaly and cyber-threat detection in a wind turbine. Google Patents. US Patent App. 15/988, 515.
- Acemoglu, D., Nedić, A., & Ozdaglar, A. (2008, December). Convergence of rule-of-thumb learning rules in social networks. In 47th IEEE conference on decision and control (pp. 1714–1720). IEEE.
- Azizi, S. M., & Khorasani, K. (2014). Networked estimation of relative sensing multiagent systems using reconfigurable sets of subobservers. *IEEE Transactions on Control Systems Technology*, 22(6), 2188–2204. https://doi.org/10.1109/TCST.2014.2300198
- Barboni, A., Rezaee, H., Boem, F., & Parisini, T. (2020). Detection of covert cyber-attacks in interconnected systems: A distributed model-based approach. *IEEE Transactions on Automatic Control*, 65(9), 3728–3741. https://doi.org/10.1109/TAC.9
- Battilotti, S., & Cacace, F. (2021). A stability with optimality analysis of consensus-based distributed filters for discrete-time linear systems. *Automatica*, 129, Article 109589. https://doi.org/10.1016/j.automatica. 2021.109589
- Battistelli, G., Chisci, L., Mugnai, G., Farina, A., & Graziano, A. (2012). Consensus-based algorithms for distributed filtering. In 51st IEEE conference on decision and control (pp. 794–799).
- Boukhobza, T., Hamelin, F., Martinez-Martinez, S., & Sauter, D. (2009, October). Structural analysis of the partial state and input observability for structured linear systems: Application to distributed systems. *European Journal of Control*, 15(5), 503–516. https://doi.org/10.3166/ejc.15. 503-516
- Charalambous, T., & Hadjicostis, C. N. (2013). Distributed formation of balanced and bistochastic weighted digraphs in multi-agent systems. In *European control conference* (pp. 1752–1757). IEEE.
- Chen, Y., Kar, S., & Moura, J. M. F. (2018). The internet of things: Secure distributed inference. *IEEE Signal Processing Magazine*, 35(5), 64–75. https://doi.org/10.1109/MSP.79
- Chong, M. S., Wakaiki, M., & Hespanha, J. P. (2015). Observability of linear systems under adversarial attacks. In *American control conference* (pp. 2439–2444). IEEE.
- Commault, C., Dion, J., & Agha, S. Y. (2008). Structural analysis for the sensor location problem in fault detection and isolation. *Automatica*, 44(8), 2074–2080. https://doi.org/10.1016/j.automatica.2007.12.014
- Commault, C., Dion, J., & Yacoub Agha, S. (2006). Location of additional sensors for FDI. In *4th workshop on advanced control and diagnosis*. Semantec Scholar.
- Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). Introduction to algorithms. MIT Press.
- Davoodi, M. R., Khorasani, K., Talebi, H. A., & Momeni, H. R. (2014). Distributed fault detection and isolation filter design for a network of heterogeneous multiagent systems. *IEEE Transactions on Control Systems Technology*, 22(3), 1061–1069. https://doi.org/10.1109/TCST.87
- Deghat, M., Ugrinovskii, V., Shames, I., & Langbort, C. (2019). Detection and mitigation of biasing attacks on distributed estimation networks. *Automatica*, 99, 369–381. https://doi.org/10.1016/j.automatica.2018. 10.052
- Dion, J. M., Commault, C., & van der Woude, J. (2003). Generic properties and control of linear structured systems: A survey. *Automatica*, 39(7), 1125–1144. https://doi.org/10.1016/S0005-1098(03)00104-3
- Doostmohammadian, M. (2019, March). Minimal driver nodes for structural controllability of large-scale dynamical systems: Node classification. *IEEE Systems Journal*, 14(3), 4209–4216. https://doi.org/10.1109/ JSYST.4267003

- Doostmohammadian, M., Charalambous, T., Shafie-khah, M., Meskin, N., & Khan, U. A. (2021). Simultaneous distributed estimation and attack detection/isolation in social networks: Structural observability, kronecker-product network, and chi-square detector. In 1st IEEE international conference on autonomous systems (ICAS) (pp. 344–348). IEEE.
- Doostmohammadian, M., & Khan, U. (2013a). On the genericity properties in distributed estimation: Topology design and sensor placement. *IEEE Journal of Selected Topics in Signal Processing*, 7(2), 195–204. https://doi.org/10.1109/JSTSP.2013.2246135
- Doostmohammadian, M., & Khan, U. A. (2013b, May). On the distributed estimation of rank-deficient dynamical systems: A generic approach. In *IEEE conference on acoustics, speech, and signal processing* (pp. 4618–4622). IEEE.
- Doostmohammadian, M., & Khan, U. A. (2016). Measurement partitioning and observational equivalence in state estimation. In *IEEE international conference on acoustics, speech and signal processing* (pp. 4855–4859). IEEE.
- Doostmohammadian, M., & Khan, U. A. (2020). Minimal sufficient conditions for structural observability/Controllability of composite networks via kronecker product. *IEEE Transactions on Signal and Information Processing Over Networks*, 6, 78–87. https://doi.org/10.1109/ TSIPN.6884276
- Doostmohammadian, M., Khan, U. A., Pirani, M., & Charalambous, T. (2021). Consensus-based distributed estimation in the presence of heterogeneous, time-Invariant delays. *IEEE Control Systems Letters*, 6. Early Access.
- Doostmohammadian, M., & Meskin, N. (2020). Sensor fault detection and isolation via networked estimation: Full-Rank dynamical systems. *IEEE Transactions on Control of Network Systems*, 8(2), 987–996. https://doi.org/10.1109/TCNS.2020.3029165
- Doostmohammadian, M., Rabiee, H. R., & Khan, U. A. (2018). Structural cost-optimal design of sensor networks for distributed estimation. *IEEE Signal Processing Letters*, 25(6), 793–797. https://doi.org/10.1109/LSP.20 18.2824761
- Doostmohammadian, M., Rabiee, H. R., & Khan, U. A. (2020). Cyber-social systems: Modeling, inference, and optimal design. *IEEE Systems Journal*, 14(1), 73–83. https://doi.org/10.1109/JSYST.4267003
- Doostmohammadian, M., Rabiee, H. R., Zarrabi, H., & Khan, U. (2018). Observational equivalence in system estimation: Contractions in complex networks. *IEEE Transactions on Network Science and Engineering*, 5(3), 212–224. https://doi.org/10.1109/TNSE.2017.2746570
- Dulmage, A., & Mendelsohn, N. (1958). Coverings of bipartite graphs. Canadian Journal of Mathematics, 10, 517–534. https://doi.org/10.4153/ CJM-1958-052-0
- Ghaoui, L. E., Oustry, F., & Rami, M. A. (1997, August). A cone complementarity linearization algorithm for static output-feedback and related problems. *IEEE Transactions on Automatic Control*, 42(8), 1171–1176. https://doi.org/10.1109/9.618250
- Giraldo, J., Urbina, D., Cardenas, A., Valente, J., Faisal, M., Ruths, J., Tippenhauer, N. O., Sandberg, H., & Candell, R. (2018). A survey of physics-based attack detection in cyber-physical systems. ACM Computing Surveys, 51(4), 1–36. https://doi.org/10.1145/3203245
- Greenwood, P. E., & Nikulin, M. S. (1996). A guide to chi-squared testing. John Wiley & Sons.
- Guan, Y., & Ge, X. (2017). Distributed attack detection and secure estimation of networked cyber-physical systems against false data injection attacks and jamming attacks. *IEEE Transactions on Signal and Information Processing Over Networks*, 4(1), 48–59. https://doi.org/10.1109/TSI PN.6884276
- Hadjicostis, C. N., & Charalambous, T. (2014). Average consensus in the presence of delays in directed graph topologies. *IEEE Transactions on Automatic Control*, 59(3), 763–768. https://doi.org/10.1109/TAC.2013. 2275669
- Harary, F. (1962, July). The determinant of the adjacency matrix of a graph. *SIAM Review*, 4(3), 202–210. https://doi.org/10.1137/1004057
- He, X., Hu, C., Hong, Y., Shi, L., & Fang, H. (2019). Distributed Kalman filters with state equality constraints: Time-based and event-triggered communications. *IEEE Transactions on Automatic Control*, 65(1), 28–43. https://doi.org/10.1109/TAC.9

- He, X., Ren, X., Sandberg, H., & Johansson, K. H. (2019). Secure distributed filtering for unstable dynamics under compromised observations. In *IEEE 58th conference on decision and control* (Vol. 67, pp. 5344–5349). IEEE.
- He, X., Ren, X., Sandberg, H., & Johansson, K. H. (2021). How to secure distributed filters under sensor attacks?. *IEEE Transactions on Automatic Control*, 67(6), arXiv preprint arXiv:2004.05409.
- Hellwig, A., & Volkmann, L. (2008). Maximally edge-connected and vertex-connected graphs and digraphs: A survey. *Discrete Mathematics*, 308(15), 3265–3296. https://doi.org/10.1016/j.disc.2007.06.035
- Ilić, M. D., Xie, L., Khan, U. A., & Moura, J. M. F. (2010). Modeling of future cyber-physical energy systems for distributed sensing and control. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 40(4), 825–838. https://doi.org/10.1109/TSMCA.2010.2048026
- Kalman, R. (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME – Journal of Basic Engineering*, 82(2), 35–45. https://doi.org/10.1115/1.3662552
- Kar, S., & Moura, J. M. F. (2011, April). Gossip and distributed Kalman filtering: Weak consensus under weak detectability. *IEEE Transactions on Signal Processing*, 59(4), 1766–1784. https://doi.org/10.1109/TSP.2010. 2100385
- Kar, S., Moura, J. M. F., & Ramanan, K. (2012). Distributed parameter estimation in sensor networks: Nonlinear observation models and imperfect communication. *IEEE Transactions on Information Theory*, 58(6), 3575–3605. https://doi.org/10.1109/TIT.2012.2191450
- Kar, S., Sinopoli, B., & Moura, J. M. F. (2011). Kalman filtering with intermittent observations: Weak convergence to a stationary distribution. *IEEE Transactions on Automatic Control*, 57(2), 405–420. https://doi.org/10.1109/TAC.2011.2161834
- Khan, U. A., & Doostmohammadian, M. (2011, December). A sensor placement and network design paradigm for future smart grids. In 4th IEEE international workshop on computational advances in multi-sensor adaptive processing (pp. 137–140). IEEE.
- Khan, U. A., & Jadbabaie, A. (2011, December). Coordinated networked estimation strategies using structured systems theory. In 49th IEEE conference on decision and control (pp. 2112–2117). IEEE.
- Khan, U. A., & Jadbabaie, A. (2014). Collaborative scalar-gain estimators for potentially unstable social dynamics with limited communication. *Automatica*, 50(7), 1909–1914. https://doi.org/10.1016/j.automatica.2014. 05.008
- Khan, U. A., & Stanković, A. M. (2013). Secure distributed estimation in cyber-physical systems. In *IEEE international conference on acoustics*, speech and signal processing (pp. 5209–5213). IEEE.
- Kodakkadan, A. R., Pourasghar, M., Puig, V., Olaru, S., Ocampo-Martinez, C., & Reppa, V. (2017). Observer-based sensor fault detectability: About robust positive invariance approach and residual sensitivity. *IFAC-PapersOnLine*, 50(1), 5041–5046. https://doi.org/10.1016/j.ifacol.2017. 08.926
- Kurt, M. N., Yılmaz, Y., & Wang, X. (2018). Distributed quickest detection of cyber-attacks in smart grid. *IEEE Transactions on Information Forensics and Security*, 13(8), 2015–2030. https://doi.org/10.1109/TIFS.2018. 2800908
- Lau, L. C., Naor, J., Salavatipour, M. R., & Singh, M. (2009). Survivable network design with degree or order constraints. SIAM Journal on Computing, 39(3), 1062–1087. https://doi.org/10.1137/070700620
- Lee, C., Shim, H., & Eun, Y. (2015). Secure and robust state estimation under sensor attacks, measurement noises, and process disturbances: Observer-based combinatorial approach. In *European control conference* (pp. 1872–1877). IEEE.
- Lee, C., Shim, H., & Eun, Y. (2019). On redundant observability: From security index to attack detection and resilient state estimation. *IEEE Transactions on Automatic Control*, 64(2), 775–782. https://doi.org /10.1109/TAC.2018.2837107
- Li, Z., & Jaimoukha, I. M. (2009). Observer-based fault detection and isolation filter design for linear time-invariant systems. *International Journal* of Control, 82(1), 171–182. https://doi.org/10.1080/00207170802031528
- Li, Z., Mazars, E., Zhang, Z., & Jaimoukha, I. M. (2012). State–space solution to the H_{-}/H_{∞} fault-detection problem. *International Journal of Robust and Nonlinear Control*, 22(3), 282–299. https://doi.org/10.1002/rnc.1690

- Lin, C. (1974, June). Structural controllability. IEEE Transactions on Automatic Control, 19(3), 201–208. https://doi.org/10.1109/TAC.1974.11 00557
- Liu, Y. Y., Slotine, J. J., & Barabási, A. L. (2011, May). Controllability of complex networks. *Nature*, 473(7346), 167–173. https://doi.org/10.1038/ nature10011
- Lopes, C. G., & Sayed, A. H. (2008, July). Diffusion least-mean squares over adaptive networks: Formulation and performance analysis. *IEEE Transactions on Signal Processing*, 56(7), 3122–3136. https://doi.org/10.1109/ TSP.2008.917383
- Micali, S., & Vazirani, V. V. (1980). An $\mathcal{O}(\sqrt{|V|}|E|)$ algorithm for finding maximum matching in general graphs. In 21st annual symposium on foundations of computer science (pp. 17–27). IEEE.
- Mitra, A., Ghawash, F., Sundaram, S., & Abbas, W. (2021). On the impacts of redundancy, diversity, and trust in resilient distributed state estimation. *IEEE Transactions on Control of Network Systems*, 8(2), 713–724. https://doi.org/10.1109/TCNS.2021.3050032
- Mitra, A., & Sundaram, S. (2018). Secure distributed state estimation of an LTI system over time-varying networks and analog erasure channels. In *American control conference* (pp. 6578–6583). IEEE.

Murota, K. (2000). Matrices and matroids for systems analysis. Springer.

- Nasrolahi, S. S., & Abdollahi, F. (2018). Sensor fault detection and recovery in satellite attitude control. *Acta Astronautica*, *145*, 275–283. https://doi.org/10.1016/j.actaastro.2018.01.002
- Nesterov, Y., & Nemirovskii, A. (1994). Interior-point polynomial algorithms in convex programming. SIAM.
- Nweke, L. O., Weldehawaryat, G. K., & Wolthusen, S. D. (2021). Threat modelling of cyber–physical systems using an applied π -Calculus. *International Journal of Critical Infrastructure Protection*, 35, Article 100466. https://doi.org/10.1016/j.ijcip.2021.100466
- Olfati-Saber, R. (2009, December). Kalman-consensus filter: Optimality, stability, and performance. In *48th IEEE conference on decision and control* (pp. 7036–7042). IEEE.
- Pajic, M., Tabuada, P., Lee, I., & Pappas, G. J. (2015). Attack-resilient state estimation in the presence of noise. In 54th IEEE conference on decision and control (pp. 5827–5832). IEEE.
- Park, S., & Martins, N. (2012). Necessary and sufficient conditions for the stabilizability of a class of LTI distributed observers. In 51st IEEE conference on decision and control (pp. 7431–7436). IEEE.
- Rahimian, M. A., & Preciado, V. M. (2015). Detection and isolation of failures in directed networks of LTI systems. *IEEE Transactions on*

Control of Network Systems, 2(2), 183–192. https://doi.org/10.1109/TC NS.2014.2378873

- Rank, M. L., & Niemann, H. (1999). Norm based design of fault detectors. International Journal of Control, 72(9), 773–783. https://doi.org/10.1080 /002071799220704
- Reinschke, K. J. (1988). Multivariable control, a graph theoretic approach. Springer.
- Ren, W., & Al-Saggaf, U. M. (2017). Distributed Kalman–Bucy filter with embedded dynamic averaging algorithm. *IEEE Systems Journal*, 12(2), 1722–1730. https://doi.org/10.1109/JSYST.2017.2657765
- Renganathan, V., Hashemi, N., Ruths, J., & Summers, T. H. (2020). Distributionally robust tuning of anomaly detectors in cyber-physical systems with stealthy attacks. In *American control conference* (pp. 1247–1252). IEEE.
- Sadeghi, E., & Fan, N. (2019). On the survivable network design problem with mixed connectivity requirements. *Annals of Operations Research*, 1–13.
- Sundaram, S. (2012). Fault-tolerant and secure control systems. University of Waterloo, Lecture Notes.
- Umsonst, D. (2019). Tuning of anomaly detectors in the presence of sensor attacks [Unpublished doctoral dissertation]. KTH Royal Institute of Technology.
- Weerakkody, S., Liu, X., & Sinopoli, B. (2017). Robust structural analysis and design of distributed control systems to prevent zero dynamics attacks. In 56th IEEE conference on decision and control (pp. 1356–1361). IEEE.
- Xiao, L., Boyd, S., & Kim, S. (2005). Distributed average consensus with least-mean-square deviation. *Journal of Parallel and Distributed Computing*, 67(1), 33–46. https://doi.org/10.1016/j.jpdc.2006.08.010
- Yan, W., Mestha, L. K., & Abbaszadeh, M. (2019). Attack detection for securing cyber physical systems. *IEEE Internet of Things Journal*, 6(5), 8471–8481. https://doi.org/10.1109/JIoT.6488907
- Ye, Y. (1993). A fully polynomial-time approximation algorithm for computing a stationary point of the general linear complementarity problem. *Mathematics of Operations Research*, 18(2), 334–345. https://doi.org/10.1287/moor.18.2.334
- Yılmaz, Y., Li, S., & Wang, X. (2016). Sequential joint detection and estimation: Optimum tests and applications. *IEEE Transactions on Signal Pro*cessing, 64(20), 5311–5326. https://doi.org/10.1109/TSP.2016.2582469
- Zhang, D., Chen, Y., Guo, F., Karimi, H. R., Dong, H., & Xuan, Q. (2021). A new interpretable learning method for fault diagnosis of rolling bearings. *IEEE Transactions on Instrumentation and Measurement*, 70, 1–10.