Sensor Fault-Tolerant State Estimation by Networks of Distributed Observers

Guitao Yang, Hamed Rezaee, *Member, IEEE*, Andrea Serrani, *Member, IEEE*, and Thomas Parisini, *Fellow, IEEE*

***Abstract*—We propose a state estimation methodology using a network of distributed observers. We consider a scenario in which the local measurement at each node may not guarantee the system's observability. In contrast, the ensemble of all the measurements does ensure that the observability property holds. As a result, we design a network of observers such that the estimated state vector computed by each observer converges to the system's state vector by using the local measurement and the communicated estimates of a subset of observers in its neighborhood. The proposed estimation scheme exploits sensor redundancy to provide robustness against faults in the sensors. Under suitable conditions on the redundant sensors, we show that it is possible to mitigate the effects of a class of sensor faults on the state estimation. Simulation trials demonstrate the effectiveness of the proposed distributed estimation scheme.**

***Index Terms*—Distributed state estimation, fault-tolerant observers, geometric-based state estimation .**

I. Introduction

THE state estimation problem with distributed measurements for a class of Lipschitz nonlinear systems is addressed in this article. Consider a scenario where measurements of a plant output are distributed in N nodes, as shown in Fig. 1. The objective is to design a distributed group of N observers located in the N nodes such that each observer provides an estimate of the whole state vector of a dynamical system. Each observer has only access to local measurements and information received from a specific subset of other nodes. A communication network with a given specific topology enables the information flow among the nodes (this is a generic *distributed estimation problem*, as formulated in [1].)

Manuscript received 31 January 2022; accepted 13 May 2022. Date of publication 13 July 2022; date of current version 27 September 2022. This work was supported in part by the European Union's Horizon 2020 research and innovation program under Grant 739551 (KIOS CoE) and in part by the Italian Ministry for Research in the framework of the 2017 Program for Research Projects of National Interest (PRIN), Grant 2017YKXYXJ. Recommended by Associate Editor A. Zemouche. *(Corresponding author: Thomas Parisini.)*

Guitao Yang and Hamed Rezaee are with the Department of Electrical and Electronic Engineering, Imperial College London, SW7 2AZ London, U.K. (e-mail: guitao.yang16@imperial.ac.uk; h.rezaee@imperial.ac.uk).

Andrea Serrani is with the Department of Electrical and Computer Engineering, The Ohio State University, Columbus, OH 43210 USA (e-mail: serrani.1@osu.edu).

Thomas Parisini is with the Department of Electrical and Electronic Engineering, Imperial College London, SW7 2AZ London, U.K., with the Department of Engineering and Architecture, University of Trieste, 34127 Trieste, Italy, and also with the KIOS Research and Innovation Center of Excellence, University of Cyprus, CY-1678 Nicosia, Cyprus (e-mail: t.parisini@gmail.com).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TAC.2022.3190429.

Digital Object Identifier 10.1109/TAC.2022.3190429

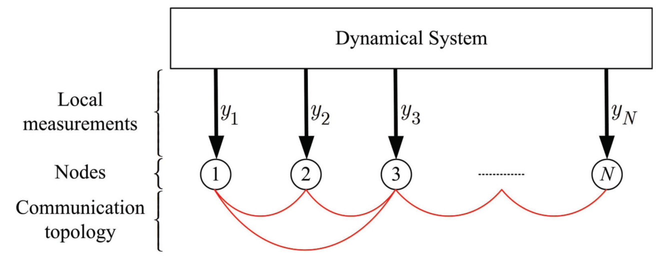


Fig. 1. Example of a network of distributed observers.

The considered state estimation problem pertains to a class of dynamical system described as

$$\dot{x} = Ax + f(x) + Bu \tag{1}$$

where $x \in \mathbb{R}^n$ is the state vector, $u \in \mathbb{R}^m$ is the control input, $A \in \mathbb{R}^{n\times n}$ is the state matrix, $B \in \mathbb{R}^{n\times m}$ is the input matrix, and $f(x) \in \mathbb{R}^n$ is a nonlinear vector field. The measurement at Node i, $i \in \mathbb{N}$, is generated according to the output equation

$$y_i = C_i x \tag{2}$$

where $y_i \in \mathbb{R}^{p_i}$ (see, for instance, [2].)

Our objective is to develop a distributed state estimation scheme that exploits sensor redundancy to tolerate possible sensor faults. We consider local sensor/output redundancy within each node, obtained by multiple measurements y_i by using a group of identical sensors at Node i. Local observers located at the nodes make up the distributed estimation scheme. Specifically, the local observer at Node i yields an estimated state vector $\hat{x}_i$ (as the estimated value of the state vector x) by using the local measurement of y_i and the estimated state vectors obtained by neighboring nodes.

To guarantee distributed observability, it is assumed that the system described by (1) and (2) is *jointly observable* [2], which means that the pair (C, A) is observable with $C = [C_1^\top\ C_2^\top\ \dots\ C_N^\top]^\top$. It should be mentioned that the pair (C_i, A) is not necessarily observable for a given Node i, and thus, there are no trivial solutions for the distributed estimation problem.

A. Brief Literature Review

Distributed state estimation has been an active area of research in the last decade as far as linear systems are concerned [1]–[13]. For instance, authors in [3]–[5] proposed distributed observer designs via the extension of the classical Kalman filter to a distributed observer network. In [6] and [7], the classical Luenberger observer is extended to a distributed estimation scheme. In [8], a distributed estimation algorithm based on a moving horizon paradigm is introduced. A linear suboptimal consensus-based distributed estimation methodology is addressed in [9]. By introducing the notion of a multisensor observable canonical decomposition, a Luenberger-based distributed linear observer design is devised in [1]. In [10], necessary and sufficient conditions for distributed estimation in linear systems are derived by introducing augmented states satisfying scalability conditions. A family of distributed observers capable of estimating state vectors with a predetermined but arbitrary rate is proposed in [2] and [11]. In [12] and [13], the problem of resilient state estimation of linear systems by increasing the robustness of communication graphs is studied.

While several works on distributed state observation have already appeared in the context of nonlinear systems [14]–[19], the methodology proposed in our article differs from the cited works as we do not assume that local observability holds. Other works deal with distributed filtering over sensor networks such that a desired stochastic performance is achieved [14]–[19]. Very recently, Chong et al. [20] proposed a secure state estimation algorithm for nonlinear systems in the presence of sensor attacks, but the observer is not designed in a distributed fashion. Finally, He et al. [21] and [22] analyzed distributed state estimation schemes for classes of jointly observable nonlinear networks. These works consider unknown nonlinearities in a discrete-time model as states to be estimated, which is distinct from our approach.

B. Objectives, Contribution, and Article Organization

The main contribution of this article is as follows.

1) We address the distributed state estimation in the presence of Lipschitz nonlinearities in the system's model. This is a step ahead from the results in [1]–[13], which are limited to linear systems.
2) We propose a robust estimation strategy against a range of additive faults in the sensors under appropriate conditions on the sensor redundancy.

More specifically, based on standard ideas in state estimation of Lipschitz nonlinear systems (for instance, see [23] and [24]), we propose a network of distributed observers for a class of Lipschitz nonlinear systems with distributed output measurements. We derive sufficient conditions on the proposed observers to guarantee that the estimated state vector of each observer converges toward the state vector of the system by using a subset of the local measurements and the estimates of the observers in its neighborhood. Moreover, in the case of sensor redundancies in each observer, we modify the proposed estimation scheme to locally reject the effects of a range of measurement faults.

Our article provides a significant generalization of the early contribution [25] in that sensors faults and disturbances are considered and sensor redundancy is exploited to mitigate the effects of the above disturbances.

The rest of this article is organized as follows: In Section II, we state the notation and provide essential notions of graph theory used in our work. We formulate the estimation problem in Section III, and present our methodology in Sections IV and V. Section VI reports simulation results showing the effectiveness of the proposed distributed fault-tolerant estimation algorithm. Finally, Section VII concludes this article.

II. Preliminaries

Here, we present notation, definitions, and basic concepts concerning graph theory.

A. Notation

Let $\mathbb{R}$ denote the set of real numbers, and $\mathbb{R}_{>0}$ and $\mathbb{R}_{\geq 0}$ denote the sets of positive and nonnegative real numbers, respectively. The set of natural numbers is denoted by $\mathbb{N}$. I_n denotes an $n \times n$ identity matrix, $\mathbf{0}_{n\times m}$ is an $n \times m$ matrix of zeros, and $\mathbf{1}_n$ is an $n \times 1$ vector of ones. $\otimes$ stands for the Kronecker product. For a matrix $A \in \mathbb{R}^{n\times m}$, $A^{-r} \in \mathbb{R}^{m\times n}$ is the right inverse of A such that $AA^{-r} = I_n$. $\|\cdot\|$ denotes the standard 2-norm. We let $\sup$ denote supremum, whereas $\min$ and $\max$ denote minimum and maximum, respectively. The function ceil$(\cdot)$ denotes the ceiling function, that is, for $x \in \mathbb{R}$, ceil(x) yields the smallest integer greater than or equal to x. For a real symmetric matrix, $\lambda_{\min}(\cdot)$ and $\lambda_{\max}(\cdot)$ denote the minimum and maximum eigenvalues, respectively, whereas $\lambda_2(\cdot)$ denotes the second smallest eigenvalue. For a square matrix M, $M \succ 0$, and $M \succeq 0$ denote positive definiteness and positive semidefiniteness, respectively. By Im and Ker, we denote, respectively, the image and the kernel of a matrix. The matrix $\mathrm{diag}(M_1, M_2, \ldots, M_n)$ is a block diagonal matrix composed of the matrices $M_1, M_2, \ldots, M_n$. Let $\mathcal{N}$ denote the class of continuous and strictly increasing functions from $[0, \infty)$ to $[0, \infty)$, and $\mathcal{K}$ denote the subset of class-$\mathcal{N}$ functions that satisfy $k(0) = 0$ [26]. For a subspace $\mathscr{V} \subseteq \mathscr{X}$ of a finite-dimensional inner-product vector space $\mathscr{X}$, $\mathscr{V}^\perp$ denotes the annihilator of $\mathscr{V}$. For $\mathscr{R}, \mathscr{S} \subseteq \mathscr{X}$, we define the subspaces $\mathscr{R} + \mathscr{S} \subseteq \mathscr{X}$ and $\mathscr{R} \cap \mathscr{S} \subseteq \mathscr{X}$ as

$$\mathscr{R} + \mathscr{S} = \{r + s : r \in \mathscr{R} \text{ and } s \in \mathscr{S}\}$$
$$\mathscr{R} \cap \mathscr{S} = \{x : x \in \mathscr{R} \text{ and } x \in \mathscr{S}\}.$$

Finally, $\oplus$ indicates the direct sum, i.e., the sum of independent subspaces, and $\simeq$ represents isomorphic relation between two vector spaces or subspaces.

B. Graph Theory

Communication among observers is described by an undirected graph $\mathcal{G} = (\mathbf{N}, \mathcal{E}, \mathcal{A})$, where $\mathbf{N} = \{1, 2, \ldots, N\}$ is a finite nonempty set of nodes of the graph (describing a set

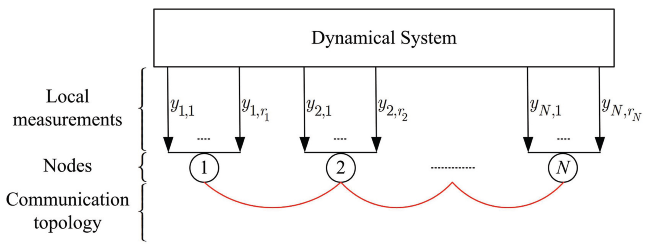


Fig. 2. Example of a network of distributed observers with measurement redundancy.

of N observers with local sensors), $\mathcal{E} \subseteq \mathbf{N} \times \mathbf{N}$ represents the edges of the graph (describing communication links), and $\mathcal{A} = [a_{ij}] \in \mathbb{R}^{N\times N}$ is the adjacency matrix, where $a_{ij} = a_{ji}$ is positive if there exists an edge between Node *i* and Node *j*, and they are zero otherwise. We define an undirected graph to be connected if there exists a path of edges containing all the nodes. The Laplacian matrix associated with the undirected graph $\mathcal{G}$ is a symmetric matrix defined as $\mathcal{L} = \mathcal{D} - \mathcal{A}$, where $\mathcal{D}$ is a diagonal matrix whose the ith entry is $d_i = \sum_{j=1}^{N} a_{ij}$. If a graph $\mathcal{G}$ is connected, the corresponding Laplacian matrix $\mathcal{L}$ has a simple zero eigenvalue with corresponding eigenvector $\mathbf{1}_N$, and all the other eigenvalues of $\mathcal{L}$ are positive real [27].

III. Problem Statement

Consider the dynamical system described in (1). We assume that observers along with their local sensors are distributed in $N \in \mathbb{N}$ nodes, and the sensor redundancy at Node i is $r_i \in \mathbb{N}$. In other words, we assume that r_i clusters of identical sensors are available at Node i. By defining $C_i \in \mathbb{R}^{p_i \times n}$ as the output matrix associated with the ith node, and by defining $\mathbf{r}_i = \{1, 2, \dots, r_i\}$, the output measurement of the kth cluster of sensors is written as follows:

$$y_{i,k} = C_i x + \varpi_{i,k}(t) + \delta_{i,k}(t), \quad i \in \mathbf{N},\ k \in \mathbf{r}_i \tag{3}$$

where vectors $\varpi_{i,k}(t) : \mathbb{R}_{\geq 0} \to \mathbb{R}^{p_i}$ and $\delta_{i,k}(t) : \mathbb{R}_{\geq 0} \to \mathbb{R}^{p_i}$, respectively, represent unknown signals denoting additive noise and faults in the kth cluster of the sensors of Node i. The matrix A is constructed such that the system is jointly observable, i.e., the pair (C, A) is observable, where $C = \begin{bmatrix} C_1^\top & C_2^\top & \cdots & C_N^\top \end{bmatrix}^\top$. Accordingly,

1) none of the pairs $(C_i, A), i \in \mathbf{N}$, may be observable, or just some of them may be observable;
2) while the pair (C, A) is observable, some nodes may even have no sensors/measurements at all.

Notice that the output redundancy is from the r_i identical local measurements defined in (3) for each C_i, whereas joint observability is related to the collection of all $C_i, i \in \mathbf{N}$. Fig. 2 shows a schematic representation of a network of distributed observers with measurement redundancy.

Assumption 1: The communication graph associated with the observers network is connected.

Assumption 2: The nonlinear vector field $f(x)$ is assumed to be globally Lipschitz, i.e., there exists $\gamma \in \mathbb{R}_{>0}$ such that

$$\|f(x_1) - f(x_2)\| \leq \gamma \|x_1 - x_2\| \quad \forall x_1, x_2 \in \mathbb{R}^n.$$

Definition 1: Decompose $y_{i,k}$, $\varpi_{i,k}$, and $\delta_{i,k}$ as

$$y_{i,k} = \begin{bmatrix} y_{i,k,1} & y_{i,k,2} & \dots & y_{i,k,p_i} \end{bmatrix}^\top$$

$$\varpi_{i,k} = \begin{bmatrix} \varpi_{i,k,1} & \varpi_{i,k,2} & \dots & \varpi_{i,k,p_i} \end{bmatrix}^\top$$

$$\delta_{i,k} = \begin{bmatrix} \delta_{i,k,1} & \delta_{i,k,2} & \dots & \delta_{i,k,p_i} \end{bmatrix}^\top$$

and let $\mathbf{p}_i = \{1, 2, \dots, p_i\}$. For a given scalar constant $\beta_{i,q} \in \mathbb{R}, i \in \mathbf{N}, q \in \mathbf{p}_i$, we define an output measurement $y_{i,k,q}, i \in \mathbf{N}, k \in \mathbf{r}_i, q \in \mathbf{p}_i$, to be *healthy* if for all $t \geq 0$

$$|\varpi_{i,k,q}(t) + \delta_{i,k,q}(t)| \leq \beta_{i,q}. \tag{4}$$

Otherwise, the measurement $y_{i,k,q}$ is called *unhealthy*. Accordingly, the associated sensors are called healthy and unhealthy, respectively. Indeed, the value of $\beta_{i,q}$ will be determined according to the required accuracy for state estimation, and this value will be considered when designing the proposed distributed observer.

Assumption 3: In the set $\{y_{i,1,q}, y_{i,2,q}, \dots, y_{i,r_i,q}\}$, at least $\mathrm{ceil}(\frac{r_i+1}{2})$ output measurements are healthy (no matter how large the magnitudes of the faults of unhealthy sensors are).

Assumption 3 implies that the number of healthy measurements at each node, i.e., those satisfying (4), exceeds that of the unhealthy ones. For example, out of five output measurements, at least three should be always healthy.

Note 1: In the rest of the article, for any variable $\vartheta_{i,k}$, $\vartheta_{i,q}$, or $\vartheta_{i,k,q}$, we let $i \in \mathbf{N}$ denote the index of the nodes, $k \in \mathbf{r}_i$ denote the index of the clusters of the sensors at Node i, and $q \in \mathbf{p}_i$ denote the index of the measurements at each cluster of the sensors at Node i.

For the nonlinear system (1) under the assumptions listed above, the objective is to design a distributed sensor fault-tolerant observer such that

1) the estimated state vectors of the observers will be the only information exchanged among the local observers at each node;
2) the estimation error is uniformly ultimately bounded.

It should be noted that the aforementioned objective also is achieved by several studies in the literature [1]–[13]. However, those studies are limited to linear systems and when available sensors are fully precise.

To achieve the objectives above, in the next section we first devise a nonlinear distributed observer scheme without considering measurement noise and faults.

IV. Distributed Observer Scheme of Jointly Observable Nonlinear Systems

The basic idea in distributed state estimation is that the ith observer updates its estimated state vector $\hat{x}_i$ not just on the basis of local measurements but also using relative information from estimated states from neighboring observers. This is typically accomplished by the introduction of an injection term of the form $\sum_{j=1}^{N} a_{ij}(\hat{x}_j - \hat{x}_i)$ (see [1], [10], [11], and [13]). Using this paradigm, the proposed distributed estimation strategy for the system described in (1) and (3) will be presented by first assuming that $\varpi_{i,k}(t) = \delta_{i,k}(t) = \mathbf{0}_{p_i\times 1}\ \forall i \in \mathbf{N} \forall k \in \mathbf{r}_i \forall t \geq 0$. Since sensor redundancy is unnecessary in this initial case, the output measurements of the system will be modeled as in (2).

To begin, define a similarity transformation matrix $T_i \in \mathbb{R}^{n\times n}, i \in \mathbf{N}$, as $T_i = [T_{io}\ T_{iu}]$, where $T_{iu} \in \mathbb{R}^{n\times v_i}$ is an orthonormal basis of the unobservable subspace of (C_i, A), $v_i \in \mathbb{N}$ is the dimension of the unobservable subspace of the pair (C_i, A), and $T_{io} \in \mathbb{R}^{n\times(n-v_i)}$ is an orthonormal basis such that $\text{Im}\, T_{io}$ is orthogonal to $\text{Im}\, T_{iu}$. According to the structure of T_i, we perform the coordinate transformation adapted to $\mathscr{X} = \text{Im}\, T_{io} \oplus \text{Im}\, T_{iu}$ as follows [28] ($\mathscr{X} \simeq \mathbb{R}^n$ is the state space of the system):

$$\begin{aligned} T_i^\top A T_i =& \begin{bmatrix} A_{io} & \mathbf{0}_{(n-v_i)\times v_i} \\ A_{ir} & A_{iu} \end{bmatrix} \\ C_i T_i =& \begin{bmatrix} C_{io} & \mathbf{0}_{p_i\times v_i} \end{bmatrix} \end{aligned} \tag{5}$$

where the pair (C_{io}, A_{io}) is observable. By using the similarity transformation matrix T_i, the local observer at Node i is designed as

$$\begin{aligned} \dot{\hat{x}}_i =& A\hat{x}_i + L_i(C_i\hat{x}_i - y_i) + f(\hat{x}_i) + Bu \\ &+ \chi P_i^{-1} \sum_{j=1}^{N} a_{ij}(\hat{x}_j - \hat{x}_i), \quad i \in \mathbf{N} \end{aligned} \tag{6}$$

where $L_i \in \mathbb{R}^{n\times p_i}$ and $P_i \in \mathbb{R}^{n\times n}$ are the observer gains computed as

$$\begin{aligned} L_i =& T_i \begin{bmatrix} P_{io}^{-1} C_{io}^\top H_i^\top \\ \mathbf{0}_{v_i\times p_i} \end{bmatrix} \\ P_i =& T_i \begin{bmatrix} P_{io} & \mathbf{0}_{(n-v_i)\times v_i} \\ \mathbf{0}_{v_i\times(n-v_i)} & I_{v_i} \end{bmatrix} T_i^\top \end{aligned} \tag{7}$$

and $P_{io} \in \mathbb{R}^{(n-v_i)\times(n-v_i)}$, $P_{io} \succ 0$, and $H_i \in \mathbb{R}^{p_i\times p_i}$ are matrices to be determined. Note that the gain χ determines the weight of the interaction term $\sum_{j=1}^{N} a_{ij}(\hat{x}_j - \hat{x}_i)$ in (6), and the term P_i^{-1} simplifies the mathematical analysis of the proposed distributed observer, as it will be clear in the sequel.

Lemma 1: Consider the jointly observable system given in (1) and (3). By letting

$$T_o = \begin{bmatrix} T_{1o} & T_{2o} & \dots & T_{No} \end{bmatrix} \tag{8}$$

one obtains

$$\text{Im}\, T_o = \mathscr{X}.$$

Proof: Since the columns of T_{iu} is an orthonormal basis of the unobservable subspace of (C_i, A), the image of T_{iu} is the kernel of the observability matrix of (C_i, A)

$$\text{Im}\, T_{iu} = \bigcap_{k=1}^{n} \text{Ker}\, C_i A^{k-1}. \tag{9}$$

Moreover, since T_{io} is such that T_i has full rank, one obtains

$$\text{Im}\, T_{iu} = (\text{Im}\, T_{io})^\perp. \tag{10}$$

According to the definition of T_o, it follows that [29, Sec. 0.12]

$$\text{Im}\, T_o = \left(\left(\sum_{i=1}^{N} \text{Im}\, T_{io}\right)^\perp\right)^\perp$$

implying that

$$\text{Im}\, T_o = \left(\bigcap_{i=1}^{N} (\text{Im}\, T_{io})^\perp\right)^\perp. \tag{11}$$

From (9)–(11), one obtains

$$\text{Im}\, T_o = \left(\bigcap_{i=1}^{N} \left(\bigcap_{k=1}^{n} \text{Ker}\, C_i A^{k-1}\right)\right)^\perp. \tag{12}$$

Since

$$\bigcap_{i=1}^{N} \left(\bigcap_{k=1}^{n} \text{Ker}\, C_i A^{k-1}\right) = \bigcap_{k=1}^{n} \left(\bigcap_{i=1}^{N} \text{Ker}\, C_i A^{k-1}\right)$$

from (12) it follows that

$$\text{Im}\, T_o = \left(\bigcap_{k=1}^{n} \text{Ker}\, (CA^{k-1})\right)^\perp. \tag{13}$$

As the pair (C, A) is observable, from (13) one obtains $\text{Im}\, T_o = \mathbf{0}_{n\times 1}^\perp = \mathscr{X}$ [29, Sec. 0.12], which completes the proof. ■

The analysis of the proposed distributed observer (6) is given in Theorem 1.

Theorem 1: Consider the noise-free and fault-free dynamical system described in (1) and (2) with the distributed observer given in (6) and (7), under Assumptions 1 and 2. The estimation error $e_i(t) = \hat{x}_i(t) - x(t)$ converges to zero for all $i \in \mathbf{N}$ if the matrices P_{io} and $H_i, i \in \mathbf{N}$, are obtained from the solution of the following LMI:

$$\begin{aligned} &\begin{bmatrix} I_{Nn} & \sqrt{\gamma}\tilde{P}^\top \\ \sqrt{\gamma}\tilde{P} & T_o(K - M)T_o^\top - \gamma N I_n \end{bmatrix} \succ 0 \\ &P_{io} \succ 0 \end{aligned} \tag{14}$$

where

$$\begin{aligned} \tilde{P} =& \begin{bmatrix} P_1 & P_2 & \dots & P_N \end{bmatrix} \\ K =& \text{diag}(K_{1o}, K_{2o}, \dots, K_{No}) \\ K_{io} =& -(A_{io}^\top P_{io} + P_{io} A_{io} + C_{io}^\top (H_i + H_i^\top) C_{io}) \\ M =& T_o^{-r} \sum_{i=1}^{N} \big(T_{iu} A_{ir} T_{io}^\top + T_{io} A_{ir}^\top T_{iu}^\top \\ &+ T_{iu}(A_{iu} + A_{iu}^\top) T_{iu}^\top\big) \left(T_o^{-r}\right)^\top. \end{aligned} \tag{15}$$

Moreover, the gain χ must be chosen such that

$$\chi > \frac{\left\|\Lambda + \gamma\bar{P} + \Lambda_P^\top Q^{-1}\Lambda_P\right\|}{2\lambda_2(\mathcal{L})} \tag{16}$$

where

$$\begin{aligned}
\Lambda =&\ \operatorname{diag}\left(\Lambda_1, \Lambda_2, \dots, \Lambda_N\right)\\
\Lambda_i =&\ (A + L_iC_i)^\top P_i + P_i(A + L_iC_i)\\
\bar{P} =&\ \operatorname{diag}\left(P_1^2 + I_n, P_2^2 + I_n, \dots, P_N^2 + I_n\right)\\
Q =&\ -\sum_{i=1}^{N}\left(\Lambda_i + \gamma(P_i^2 + I_n)\right)\\
\Lambda_P =&\ \left[\Lambda_1 + \gamma(P_1^2 + I_n) \quad \dots \quad \Lambda_N + \gamma(P_N^2 + I_n)\right].
\end{aligned} \tag{17}$$

Proof: To begin, from (1), (3), and (6), it follows that the dynamics of $e_i = \hat{x}_i - x$, $i \in \mathbf{N}$, are given by

$$\begin{aligned}
\dot{e}_i =&\ (A + L_iC_i)e_i + f(\hat{x}_i) - f(x)\\
&+ \chi P_i^{-1}\sum_{j=1}^{N} a_{ij}(\hat{x}_j - \hat{x}_i), \quad i \in \mathbf{N}
\end{aligned}$$

which, since $\hat{x}_j - \hat{x}_i = e_j - e_i$, can be restated as follows:

$$\begin{aligned}
\dot{e}_i =&\ (A + L_iC_i)e_i + f(\hat{x}_i) - f(x)\\
&+ \chi P_i^{-1}\sum_{j=1}^{N} a_{ij}(e_j - e_i), \quad i \in \mathbf{N}.
\end{aligned} \tag{18}$$

We shall show that the solutions of (18) converge to zero. To this end, consider the following Lyapunov candidate:

$$V = \sum_{i=1}^{N} e_i^\top P_i e_i \tag{19}$$

which is a positive definite function of the estimation error, since $P_{io} \succ 0, i \in \mathbf{N}$. The derivative of V along (18) reads as

$$\begin{aligned}
\dot{V} =&\ \sum_{i=1}^{N} e_i^\top\left((A + L_iC_i)^\top P_i + P_i(A + C_iL_i)\right)e_i\\
&+ 2\chi\sum_{i=1}^{N}\sum_{j=1}^{N} a_{ij}\left(e_j - e_i\right)^\top e_i\\
&+ 2\sum_{i=1}^{N}\left(f(\hat{x}_i) - f(x)\right)^\top P_ie_i.
\end{aligned} \tag{20}$$

According to Assumption 2, one obtains

$$\sum_{i=1}^{N}\left(f(\hat{x}_i) - f(x)\right)^\top P_ie_i \le \sum_{i=1}^{N}\gamma\|\hat{x}_i - x\|\|P_ie_i\|$$

and since $e_i := \hat{x}_i - x$, it follows that

$$2\sum_{i=1}^{N}\left(f(\hat{x}_i) - f(x)\right)^\top P_ie_i \le 2\gamma\sum_{i=1}^{N}\|e_i\|\|P_ie_i\|. \tag{21}$$

As $2\|e_i\|\|P_ie_i\| \le e_i^\top e_i + e_i^\top P_i^2 e_i$, using the definition of $\bar{P}$ in (17), one obtains

$$2\sum_{i=1}^{N}\left(f(\hat{x}_i) - f(x)\right)^\top P_ie_i \le \gamma e^\top\bar{P}e \tag{22}$$

where $e = \begin{bmatrix} e_1^\top & e_2^\top & \dots & e_N^\top\end{bmatrix}^\top$. Moreover,

$$2\chi\sum_{i=1}^{N}\sum_{j=1}^{N} a_{ij}\left(e_j - e_i\right)^\top e_i = -2\chi e^\top\left(\mathcal{L}\otimes I_n\right)e. \tag{23}$$

From (20), (22), and (23), and according to the definition of Λ in (17), it follows that

$$\dot{V} \le e^\top(\Lambda + \gamma\bar{P})e - 2\chi e^\top\left(\mathcal{L}\otimes I_n\right)e. \tag{24}$$

Since the pair (C_i, A) may not be observable, Λ may not be negative definite. Thus, to show that the right-hand side of (24) is negative definite, we decompose the error space into two complementary subspaces such that $e^\top(\Lambda + \gamma\bar{P})e$ is negative definite when projected onto one of the subspaces, and $-2\chi e^\top(\mathcal{L}\otimes I_n)e$ is negative definite when projected onto the other one. According to Assumption 1 and the properties of the Laplacian matrix of a connected undirected graph, $\mathcal{L}$ has a zero eigenvalue and $N-1$ positive real eigenvalues, and the right eigenvector associated with the zero eigenvalue is $\mathbf{1}_N$. Thus, we decompose the error e as

$$e = e_c + e_r \tag{25}$$

where $e_c \in \mathbb{R}^{Nn}$ is the *consensus vector* in the form $e_c = \mathbf{1}_N \otimes \omega$, $\omega \in \mathbb{R}^n$, and $e_r \in \mathbb{R}^{Nn}$ is the *disagreement vector* satisfying $e_r^\top e_c = 0$ [30]. The intuition behind the aforementioned decomposition is to exploit the properties of the Laplacian matrix associated with a connected graph. Since $e_c = \mathbf{1}_N \otimes \omega$, the vector e_c lies in the kernel of $\mathcal{L}\otimes I_n$, which simplifies the mathematical analysis of (24). Moreover, we will show that $e_c^\top(\Lambda + \gamma\bar{P})e_c$ is negative definite with respect to e_c. Since e_r is orthogonal to the kernel of $\mathcal{L}\otimes I_n$ and the second least eigenvalue of the Laplacian matrix is positive real, $-2\chi e_r^\top(\mathcal{L}\otimes I_n)e_r$ is negative definite with respect to e_r. Hence, using (25), the inequality (24) yields

$$\begin{aligned}
\dot{V} \le&\ e_c^\top(\Lambda + \gamma\bar{P})e_c + 2e_r^\top(\Lambda + \gamma\bar{P})e_c\\
&+ e_r^\top(\Lambda + \gamma\bar{P})e_r - 2\chi e_r^\top(\mathcal{L}\otimes I_n)e_r.
\end{aligned} \tag{26}$$

To guarantee negative definiteness of $\dot{V}$, one must show that $e_c^\top(\Lambda + \gamma\bar{P})e_c$ is negative definite. Since $e_c = \mathbf{1}_N \otimes \omega$, according to the definition of Λ and $\bar{P}$ in (17), one obtains

$$e_c^\top(\Lambda + \gamma\bar{P})e_c = \omega^\top\left(\sum_{i=1}^{N}\left(\Lambda_i + \gamma(P_i^2 + I_n)\right)\right)\omega. \tag{27}$$

Next, we show that the right-hand side of (27) is negative definite. From (5), it follows that

$$\begin{aligned}
A =&\ T_i\begin{bmatrix} A_{io} & \mathbf{0}_{(n-v_i)\times v_i}\\ A_{ir} & A_{iu}\end{bmatrix}T_i^\top\\
C_i =&\ \begin{bmatrix} C_{io} & \mathbf{0}_{p_i\times v_i}\end{bmatrix}T_i^\top.
\end{aligned} \tag{28}$$

Therefore, using the definition of L_i and P_i in (7), $\Lambda_i := (A+L_iC_i)^\top P_i + P_i(A+L_iC_i)$ can be expressed as

$$\Lambda_i = T_i \begin{bmatrix} -K_{io} & A_{ir}^\top \\ A_{ir} & A_{iu}+A_{iu}^\top \end{bmatrix} T_i^\top. \tag{29}$$

Using $T_i = \begin{bmatrix} T_{io} & T_{iu} \end{bmatrix}$, (29) can be rewritten as

$$\begin{aligned}\Lambda_i = & -T_{io}K_{io}T_{io}^\top + T_{iu}A_{ir}T_{io}^\top + T_{io}A_{ir}^\top T_{iu}^\top \\ & + T_{iu}(A_{iu}+A_{iu}^\top)T_{iu}^\top.\end{aligned}$$

Therefore,

$$\begin{aligned}\sum_{i=1}^N \Lambda_i = & -T_oKT_o^\top + \sum_{i=1}^N \left(T_{iu}A_{ir}T_{io}^\top + T_{io}A_{ir}^\top T_{iu}^\top\right) \\ & + \sum_{i=1}^N T_{iu}(A_{iu}+A_{iu}^\top)T_{iu}^\top \end{aligned} \tag{30}$$

where $K \in \mathbb{R}^{(Nn-\sum_{i=1}^N v_i)\times(Nn-\sum_{i=1}^N v_i)}$ is the block diagonal matrix given in (15) and T_o is defined in (8). By invoking Lemma 1, T_o is full rank and hence its rows are linearly independent. Therefore, there exists T_o^{-r} such that $T_oT_o^{-r} = I_n$. Accordingly, (30) can be written in the following form:

$$\sum_{i=1}^N \Lambda_i = -T_o\,(K-M)\,T_o^\top$$

where M is defined in (15). Therefore,

$$\begin{aligned} & -\sum_{i=1}^N \left(\Lambda_i + \gamma(P_i^2+I_n)\right) \\ & = T_o\,(K-M)\,T_o^\top - \gamma N I_n - \gamma\tilde{P}\tilde{P}^\top \end{aligned} \tag{31}$$

where $\tilde{P}$ is as defined in (15). According to (14) and the Schur complement decomposition [31], one obtains

$$T_o\,(K-M)\,T_o^\top - \gamma N I_n - \gamma\tilde{P}\tilde{P}^\top \succ 0. \tag{32}$$

From (31) and (32), it follows that the right-hand side of (27) is negative definite, and as a result

$$\begin{aligned} & e_c^\top(\Lambda+\gamma\bar{P})e_c \\ & = \omega^\top \left(\sum_{i=1}^N \left(\Lambda_i + \gamma(P_i^2+I_n)\right)\right)\omega < 0 \;\forall \omega \neq \mathbf{0}_{n\times 1}. \end{aligned} \tag{33}$$

Next, we investigate the other terms in the right-hand side of the inequality (26). As $e_c = \mathbf{1}_N \otimes \omega$, one has

$$e_r^\top(\Lambda+\gamma\bar{P})e_c = e_r^\top\Lambda_P^\top\omega = \omega^\top\Lambda_P e_r \tag{34}$$

where Λ_P is defined in (17). Moreover, since $e_r^\top e_c = 0$, it follows that

$$e_r^\top(\mathbf{1}_N\otimes\omega) = 0 \quad \forall \omega \in \mathbb{R}^n.$$

Since $\mathbf{1}_N$ is the eigenvector associated with the zero eigenvalue of $\mathcal{L}$, one obtains [30]

$$-e_r^\top(\mathcal{L}\otimes I_n)e_r \le -\lambda_2(\mathcal{L})e_r^\top e_r. \tag{35}$$

According to (33), (34), and (35), (26) is satisfied if

$$\begin{aligned} & \dot{V} \le \\ & -\begin{bmatrix}\omega \\ e_r\end{bmatrix}^\top \begin{bmatrix} Q & -\Lambda_P \\ -\Lambda_P^\top & 2\chi\lambda_2(\mathcal{L})I_{Nn} - \left(\Lambda+\gamma\bar{P}\right)\end{bmatrix} \begin{bmatrix}\omega \\ e_r\end{bmatrix} \end{aligned} \tag{36}$$

where $Q \in \mathbb{R}^{n\times n} \succ 0$ is defined in (17). From (16), it follows that

$$2\chi\lambda_2(\mathcal{L})I_{Nn} - \Lambda - \gamma\bar{P} - \Lambda_P^\top Q^{-1}\Lambda_P \succ 0,$$

thus, by using Schur complements, one obtains

$$\begin{bmatrix} Q & -\Lambda_P \\ -\Lambda_P^\top & 2\chi\lambda_2(\mathcal{L})I_{Nn} - \left(\Lambda+\gamma\bar{P}\right)\end{bmatrix} \succ 0. \tag{37}$$

Consequently, letting

$$k := \lambda_{\min}\left(\begin{bmatrix} Q & -\Lambda_P \\ -\Lambda_P^\top & 2\chi\lambda_2(\mathcal{L})I_{Nn} - \left(\Lambda+\gamma\bar{P}\right)\end{bmatrix}\right) > 0 \tag{38}$$

one obtains

$$\dot{V} \le -k\left\|\begin{bmatrix}\omega \\ e_r\end{bmatrix}\right\|^2 \tag{39}$$

Since $e = e_r + e_c$ and $e_c = \mathbf{1}_N \otimes \omega$, it follows that

$$\|e_i\| \le \left\|\begin{bmatrix}\omega \\ e_r\end{bmatrix}\right\|$$

hence

$$\|e\| \le \sqrt{N}\left\|\begin{bmatrix}\omega \\ e_r\end{bmatrix}\right\|. \tag{40}$$

Inequalities (39) and (40) yield

$$\dot{V} \le -\frac{k}{\sqrt{N}}\|e\|^2. \tag{41}$$

Moreover, from (19), it follows that

$$V \le \max_{i\in\mathbf{N}}(\lambda_{\max}(P_i))\|e\|^2. \tag{42}$$

According to (41) and (42), one obtains

$$\dot{V} \le -\frac{k}{\sqrt{N}\max_{i\in\mathbf{N}}(\lambda_{\max}(P_i))}V.$$

Thus, $V(t)$ asymptotically converges to zero along solutions of (18), implying that the estimation errors $e_i(t)$ converge to zero for all $i \in \mathbf{N}$. ■

It is worth noting that, from a complexity perspective, the observer proposed in (6) has the same dimension as those introduced in [1], [3], [11], [28], and [32]. Moreover, compared to the introduced distributed estimation schemes in [10] and [2], the proposed observer does not require any augmented states.

Remark 1: It is worth mentioning that if $\gamma = 0$ (this happens when $f(x) = \mathbf{0}_{n\times 1}$,) the LMI condition (14) will be as follows:

$$\begin{bmatrix} I_{Nn} & \mathbf{0}_{Nn\times n} \\ \mathbf{0}_{n\times Nn} & T_o\,(K-M)\,T_o^\top \end{bmatrix} \succ 0. \tag{43}$$

Since (C_{io}, A_{io}) is observable for all $i \in \mathbf{N}$, for any symmetric positive definite matrix K_{io}, there exists H_i such that the Riccati-like equation $K_{io} = -(A_{io}^\top P_{io} + P_{io}A_{io} + C_{io}^\top(H_i + H_i^\top)C_{io})$ has a symmetric positive definite solution P_{io}. If this is the case, (43) is solvable if the following inequality is satisfied:

$$\lambda_{\min}(K) = \min_{i\in\mathbf{N}} \lambda_{\min}(K_{io}) > \|M\|. \tag{44}$$

Thus, by designing K based on (44), the distributed estimation problem is always feasible when $\gamma = 0$. It should be noted that the solvability of the LMI (14) is a sufficient condition in Theorem 1 to guarantee state estimation by the network of observers. However, since the distributed estimation problem of Theorem 1 is always feasible when $\gamma = 0$, there exists a set of values for γ such that the LMI (14) has solutions for P_{io} and $H_i, i \in \mathbf{N}$. Indeed, since (C_{io}, A_{io}) is observable for all $i \in \mathbf{N}$, there exist P_{io} and $H_i, i \in \mathbf{N}$, such that the following matrix has positive real eigenvalues:

$$M_1 = \begin{bmatrix} I_{Nn} & \mathbf{0}_{Nn\times n} \\ \mathbf{0}_{n\times Nn} & T_o(K-M)T_o^\top \end{bmatrix}.$$

If γ is sufficiently small, by continuity, the eigenvalues of the matrix

$$M_2 = \begin{bmatrix} I_{Nn} & \sqrt{\gamma}\tilde{P}^\top \\ \sqrt{\gamma}\tilde{P} & T_o(K-M)T_o^\top - \gamma N I_n \end{bmatrix}$$

remain positive real, implying that the LMI (14) is solvable.

The basis of Theorem 1 is the availability of precise measurements at all nodes. In the next section, the proposed strategy will be generalized to the more realistic case of measurement noise and sensor faults.

V. Sensor Fault-Tolerant Distributed State Estimation Scheme

In this section, we present the fault-tolerant and noise-resilient version of the distributed state observer developed previously. To tolerate the effect of sensor faults in the state estimation, we use redundant sensors and an appropriate selection of the measurements. The main idea of using sensor redundancy for state estimation has been the detection of the faulty sensors and the reconfiguration of the observer after the occurrence and detection of the faults [33]–[38]. There exists an inevitable latency time required to perform fault detection and the subsequent observer reconfiguration, which may not be efficient in some practical applications, especially for fast varying state systems. In this section, by considering all the sensor measurements altogether, we show that it is possible to design an observer that can reject the effect of sensor faults.

Owing to Definition 1, an output measurement $y_{i,k,q}(t)$ $i \in \mathbf{N}, k \in \mathbf{r}_i, q \in \mathbf{p}_i$ is deemed *healthy* if the norm of the corresponding measurement error is smaller than a given level, $\beta_{i,q} > 0$. Decomposing $C_i x(t)$ as

$$C_i x(t) = \begin{bmatrix} \zeta_{i,1}(t) & \zeta_{i,2}(t) & \ldots & \zeta_{i,p_i}(t) \end{bmatrix}^\top$$

an output measurement $y_{i,k,q}$ is healthy if

$$|y_{i,k,q}(t) - \zeta_{i,q}(t)| \le \beta_{i,q}\ \forall t \ge 0. \tag{45}$$

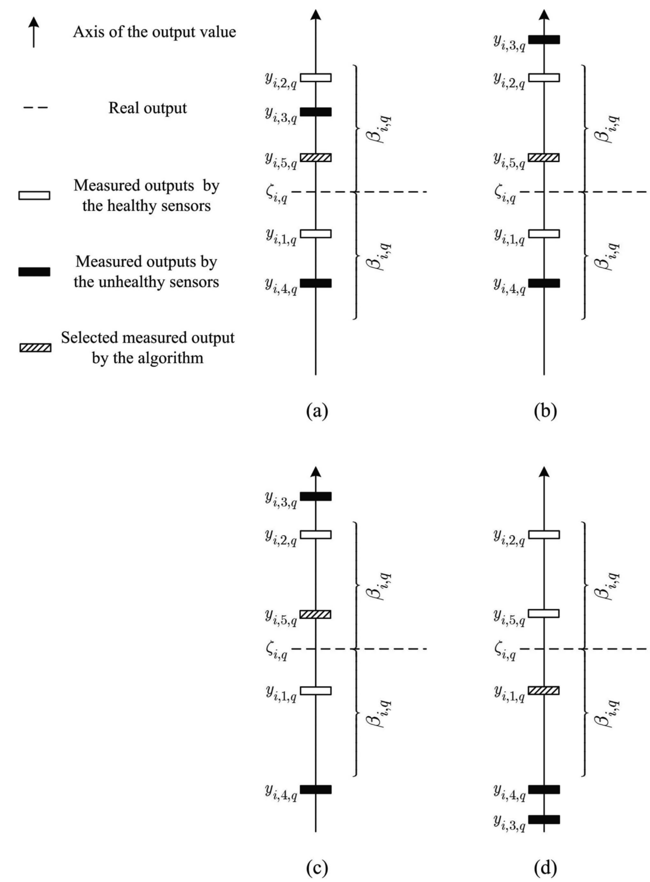


Fig. 3. Various configurations of the measured outputs of healthy and unhealthy sensors when $r_i = 5$.

Algorithm 1: Selection of Measurements.

1: For each $q \in \mathbf{p}_i$ and Node i, obtain and hold $u(t)$ and $y_{i,k,q}(t), k \in \mathbf{r}_i$.
2: Sort $y_{i,k,q}(t), k \in \mathbf{r}_i$, from the largest to the smallest one, and select $y_{i,m,q}(t)$, where $m = \text{ceil}(\frac{r_i}{2})$.
3: Return $y_{i,m,q}(t), q \in \mathbf{p}_i$, and $u(t)$.

By exploiting Assumption 3 for each $q \in \mathbf{p}_i$, Algorithm 1 detailed below makes it possible for Node i to ignore faulty output measurements and employ output measurements for which (45) is satisfied. In practice, the algorithm causes a delay which is accounted for by assuming that the observers use the information of the output measurement and the control input with an unknown constant time delay $\tau \in \mathbb{R}_{>0}$. It is worth mentioning that increasing r_i may lead to higher computational demands, hence to an increment in such delay.

To explain the algorithm with an example, consider a case where $r_i = 5$. Based on Assumption 3, at each time instant, at most two output measurements can fall outside the bound described in (45). Hence, in terms of the values of output measurements of unhealthy sensors, at each time instant four cases need to be considered, as shown in Fig. 3. In the first case in Fig. 3(a), the output measurements of unhealthy sensors are in the domain described by the bound in (45). In the second case shown in Fig. 3(b), one of them is outside the domain. In

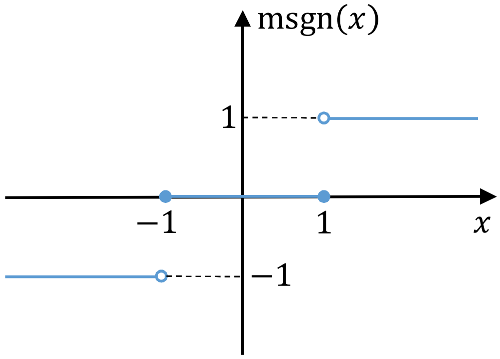


Fig. 4. msgn$(\cdot)$ function.

the remaining cases, shown in Fig. 3(c) and (d), respectively, two measurements are outside the healthy range. In all cases, only one output measurement within the range (45) will be selected by Algorithm 1. Note that according to the first two cases, in some time instants, the chosen output measurement can be the output measurement of an unhealthy sensor whose measurement is in the domain (45). Due to the abovementioned issues, Assumption 3 is needed to guarantee that Algorithm 1 can always choose measurements satisfying (4). This assumption is reasonable and practical in real-world applications when fusion of redundant sensors is needed (for instance, see [12] and [13].)

By employing the output measurement chosen by Algorithm 1 and by considering the delay due to the computation of the algorithm, we will provide a design of the distributed observer such that

$$\lim_{t\to\infty}\sup\sum_{i=1}^{N}\|\hat{x}_i(t)-x(t-\tau)\|\leq\rho\left(\sum_{i,q}\beta_{i,q}^2\right)$$

where $\rho(\cdot)\in[0,\infty)$ is a class-$\mathcal{K}$ function of the bounds $\beta_{i,q}\in\mathbb{R}$, $i\in\mathbf{N}, q\in\mathbf{p}_i$. Robustification against sensor disturbance is achieved via the nonlinear function depicted in Fig. 4, which is denoted by msgn$(\cdot)$. By tuning the dead zone domain of the function msgn$(\cdot)$, it is possible to tune the sensitivity of the function to a range of output measurement errors. Moreover, tuning the function's gain affects robustness against a range of measurement faults and noises when the output measurement error lies outside the dead zone domain. We will study these features in more detail in the sequel.

Remark 2: According to Assumption 3, for the measurement set $\{y_{i,1,q},y_{i,2,q},\ldots,y_{i,r_i,q}\}$, at least ceil$(\frac{r_i+1}{2})$ output measurements are expected to be healthy, that is, the number of healthy output measurements should be larger than the number of unhealthy ones. Therefore, in the presence of faults, r_i should be at least equal to 3 for Assumption 3 to be satisfied. Consequently, increasing r_i increases the possibility for Algorithm 1 to choose measurements satisfying (4). However, increasing r_i requires more sensors, leading to higher implementation costs.

For the system described in (1) and (3), the proposed fault-tolerant distributed observer reads as

$$\begin{aligned}\dot{\hat{x}}_i(t)=\ &A\hat{x}_i(t)+L_i\left(C_i\hat{x}_i(t)-y_{i,m}(t-\tau)-\nu_i(t)\right)\\&+f(\hat{x}_i(t))+\chi P_i^{-1}\sum_{j=1}^{N}a_{ij}(\hat{x}_j(t)-\hat{x}_i(t))\\&+Bu(t-\tau),\quad i\in\mathbf{N}_i\end{aligned}\tag{46}$$

where L_i and P_i are the same as (7) and χ is the same as (16). In addition, $\nu_i(t)$ is a robustifying term designed as follows:

$$\nu_i(t)=\boldsymbol{\eta_i}\mathrm{msgn}\big(\boldsymbol{\beta_i}^{-1}\mathrm{sgn}(H_i)(C_i\hat{x}_i(t)-y_{i,m}(t-\tau))\big)\tag{47}$$

where $\boldsymbol{\beta_i}\in\mathbb{R}^{p_i\times p_i}$, $\boldsymbol{\eta_i}\in\mathbb{R}^{p_i\times p_i}$, and $H_i\in\mathbb{R}^{p_i\times p_i}$ are diagonal matrices defined as

$$\begin{aligned}\boldsymbol{\beta_i}&=\mathrm{diag}(\beta_{i,1},\beta_{i,2},\ldots,\beta_{i,p_i})\\\boldsymbol{\eta_i}&=\mathrm{diag}(\eta_{i,1},\eta_{i,2},\ldots,\eta_{i,p_i})\\H_i&=\mathrm{diag}(H_{i,1},H_{i,2},\ldots,H_{i,p_i}).\end{aligned}$$

The entries of $\boldsymbol{\eta_i}$ and H_i will be determined in the upcoming analysis. Note that since Algorithm 1 is used for obtaining $y_{i,m}$, in (46) and (47), the time delay τ is considered in $y_{i,m}$. Hence, instead of pursuing estimates of the real-time state vector $x(t)$, we intend to estimate the delayed state $x(t-\tau)$. Consequently, in the observer design, the input u in (46) is delayed by τ via the algorithm.

Remark 3: It should be noted that (46) has a discontinuous right-hand side due to the msgn$(\cdot)$ function and Algorithm 1. However, since msgn$(\cdot)$ is locally bounded and measurable, a solution in the sense of Filippov exists for (46) [39]–[41] (specifically, see [40, Sec. 2].)

Theorem 2: Consider the dynamical system described in (1) and (3) under the distributed observer given in (46) and (47), when Assumptions 1, 2, and 3 are satisfied, L_i and P_i are the same as (7), the symmetric positive definite matrices P_{io} and the diagonal matrices $H_i, i\in\mathbf{N}$, are obtained from the solution of the LMI (14), χ is chosen as (16), and $\eta_{i,q}\geq\beta_{i,q}, i\in\mathbf{N}, q\in\mathbf{p}_i$, where $\tilde{P}$, K, M, Λ, $\bar{P}$, Q, and Λ_P are the same as (15) and (17). Under these conditions, the estimation errors $e_i(t)=\hat{x}_i(t)-x(t-\tau)$, $i\in\mathbf{N}$, $t\geq\tau$, are uniformly ultimately bounded and satisfy

$$\lim_{t\to\infty}\sup\|e(t)\|\leq 2\sqrt{\frac{\sum_{i=1}^{N}\sum_{q=1}^{p_i}|H_{i,q}|\beta_{i,q}^2}{\mu\min\limits_{i\in\mathbf{N}}(\lambda_{\min}(P_i))}}\tag{48}$$

where

$$e(t)=\begin{bmatrix}e_1(t)^\top & e_2(t)^\top & \cdots & e_N(t)^\top\end{bmatrix}^\top$$

$$\mu=\frac{k}{\sqrt{N}\max\limits_{i\in\mathbf{N}}(\lambda_{\max}(P_i))}$$

and k is defined in (38).

Proof: From (1), it follows that

$$\dot{x}(t-\tau)=Ax(t-\tau)+f(x(t-\tau))+Bu(t-\tau).\tag{49}$$

Since $\hat{x}_j(t)-\hat{x}_i(t)=e_j(t)-e_i(t)$, from (3), (46), (47), and (49), owing to Algorithm 1, the differential equation describing

$e_i(t) = \hat{x}_i(t) - x(t-\tau), i \in \mathbf{N}$, is given by

$$\begin{aligned}\dot{e}_i(t) = &\ (A + L_iC_i)e_i(t) + f(\hat{x}_i(t)) - f(x(t-\tau)) \\ &+ \chi P_i^{-1}\sum_{j=1}^{N} a_{ij}(e_j(t) - e_i(t)) \\ &- L_i\left(\varpi_{i,m}(t-\tau) + \delta_{i,m}(t-\tau)\right) \\ &- L_i\boldsymbol{\eta_i}\text{msgn}\left(\boldsymbol{\beta_i}^{-1}\text{sgn}(H_i)(C_ie_i(t)\right. \\ &\left.- \varpi_{i,m}(t-\tau) - \delta_{i,m}(t-\tau))\right). \end{aligned} \tag{50}$$

To analyze the evolution of $e_i(t)$ along (50), we consider the same Lyapunov candidate as (19). Thus, following a procedure similar to that in the proof of Theorem 1, the derivative of V along (50) reads as

$$\begin{aligned}\dot{V}(t) \leq &\ e^\top(t)(\Lambda + \gamma\bar{P})e(t) - 2\chi e^\top(t)(\mathcal{L}\otimes I_n)e(t) \\ &- \sum_{i=1}^{N} 2e_i^\top P_iL_i\left(\varpi_{i,m}(t-\tau) + \delta_{i,m}(t-\tau)\right. \\ &+ \boldsymbol{\eta_i}\text{msgn}\left(\boldsymbol{\beta_i}^{-1}\text{sgn}(H_i)(C_ie_i(t)\right. \\ &\left.\left.- \varpi_{i,m}(t-\tau) - \delta_{i,m}(t-\tau))\right)\right). \end{aligned} \tag{51}$$

Due to the structure of P_i and L_i given in (7) and since $T_i^\top T_i = I_n$, one obtains

$$P_iL_i = T_i\begin{bmatrix} P_{io} & \mathbf{0}_{(n-v_i)\times v_i} \\ \mathbf{0}_{v_i\times(n-v_i)} & I_{v_i}\end{bmatrix}\begin{bmatrix} P_{io}^{-1}C_{io}^\top H_i \\ \mathbf{0}_{v_i\times p_i}\end{bmatrix}$$

which can be simplified as follows:

$$P_iL_i = T_i\begin{bmatrix} C_{io}^\top \\ \mathbf{0}_{v_i\times p_i}\end{bmatrix}H_i. \tag{52}$$

Since $T_i\begin{bmatrix} C_{io} & \mathbf{0}_{p_i\times v_i}\end{bmatrix}^\top = C_i^\top$ and H_i is diagonal, from (52), one obtains

$$P_iL_i = C_i^\top H_i^\top. \tag{53}$$

Following a procedure similar to that in the proof of Theorem 1 for (24) and by considering (53), from (51), one obtains

$$\begin{aligned}\dot{V}(t) \leq &\ -\mu V(t) - 2\sum_{i=1}^{N}(H_iC_ie_i(t))^\top \\ &\left(\varpi_{i,m}(t-\tau) + \delta_{i,m}(t-\tau)\right. \\ &+ \boldsymbol{\eta_i}\text{msgn}\left(\boldsymbol{\beta_i}^{-1}\text{sgn}(H_i)(C_ie_i(t)\right. \\ &\left.\left.- \varpi_{i,m}(t-\tau) - \delta_{i,m}(t-\tau))\right)\right). \end{aligned} \tag{54}$$

Next, we analyze the other term in the right-hand side of (54) that is

$$\text{msgn}\left(\boldsymbol{\beta_i}^{-1}\text{sgn}(H_i)(C_ie_i(t) - \varpi_{i,m}(t-\tau) - \delta_{i,m}(t-\tau))\right).$$

According to Assumption 3 and Algorithm 1, for each $i \in \mathbf{N}$ and $q \in \mathbf{p}_i$, (4) is satisfied for the selected output measurements. By decomposing $C_ie_i(t)$ as

$$C_ie_i(t) = \begin{bmatrix}\xi_{i,1}(t) & \xi_{i,2}(t) & \dots & \xi_{i,p_i}(t)\end{bmatrix}^\top$$

for each $i \in \mathbf{N}$ and $q \in \mathbf{p}_i$, two cases arise. One case is when $|\xi_{i,q}(t)| > 2\beta_{i,q}$ and the other one is when $|\xi_{i,q}(t)| \leq 2\beta_{i,q}$. If $|\xi_{i,q}(t)| > 2\beta_{i,q}$, according to the definition of the $\text{msgn}(\cdot)$ function, it can be concluded that

$$\begin{aligned}&\text{msgn}\left(\beta_{i,q}^{-1}\text{sgn}(H_{i,q})(\xi_{i,q}(t) - \varpi_{i,m,q}(t-\tau)\right. \\ &\left.- \delta_{i,m,q}(t-\tau))\right) = \text{sgn}(H_{i,q}\xi_{i,q}(t)).\end{aligned}$$

According to (4) and since $\eta_{i,q} \geq \beta_{i,q}, i \in \mathbf{N}, q \in \mathbf{p}_i$, there is a nonnegative real scalar $\alpha_{i,q}$ such that

$$\begin{aligned}&- H_{i,q}\xi_{i,q}(t)\left(\varpi_{i,m,q}(t-\tau) + \delta_{i,m,q}(t-\tau)\right. \\ &+ \eta_{i,q}(t)\text{msgn}\left(\beta_{i,q}^{-1}\text{sgn}(H_{i,q})(\xi_{i,q}(t)\right. \\ &\left.\left.- \varpi_{i,m,q}(t-\tau) - \delta_{i,m,q}(t-\tau))\right)\right) \\ &= -H_{i,q}\xi_{i,q}(t)\alpha_{i,q}\text{sgn}(H_{i,q}\xi_{i,q}(t)).\end{aligned} \tag{55}$$

Conversely, if $|\xi_{i,q}(t)| \leq 2\beta_{i,q}$, according to the definition of the $\text{msgn}(\cdot)$ function, the measurement errors satisfying (4) cannot invert the sign of

$$\text{msgn}\left(\beta_{i,q}^{-1}\text{sgn}(H_{i,q})(\xi_{i,q}(t) - \varpi_{i,m}(t-\tau) - \delta_{i,m}(t-\tau))\right).$$

Hence, (55) is still satisfied or

$$\begin{aligned}&\text{msgn}\left(\beta_{i,q}^{-1}\text{sgn}(H_{i,q})(\xi_{i,q}(t) - \varpi_{i,m,q}(t-\tau)\right. \\ &\left.- \delta_{i,m,q}(t-\tau))\right) = 0\end{aligned}$$

which happens when

$$|\xi_{i,q}(t) - \varpi_{i,m,q}(t-\tau) - \delta_{i,m,q}(t-\tau)| \leq \beta_{i,q}.$$

If this is the case, according to (4) and the fact that $|\xi_{i,q}(t)| \leq 2\beta_{i,q}$, one obtains in place of (55)

$$\begin{aligned}&- H_{i,q}\xi_{i,q}(t)\left(\varpi_{i,m,q}(t-\tau) + \delta_{i,m,q}(t-\tau)\right. \\ &+ \eta_{i,q}\text{msgn}\left(\beta_{i,q}^{-1}\text{sgn}(H_{i,q})(\xi_{i,q}(t)\right. \\ &\left.\left.- \varpi_{i,m,q}(t-\tau) - \delta_{i,m,q}(t-\tau))\right)\right) \\ &\leq -H_{i,q}\xi_{i,q}(t)(\varpi_{i,m,q}(t-\tau) + \delta_{i,m,q}(t-\tau)) \\ &\leq 2|H_{i,q}|\beta_{i,q}^2.\end{aligned} \tag{56}$$

From (54)–(56), it follows that

$$\dot{V}(t) \leq -\mu V(t) + 4\sum_{i=1}^{N}\sum_{q=1}^{p_i}|H_{i,q}|\beta_{i,q}^2 \tag{57}$$

and since

$$V(t) \geq \min_{i\in\mathbf{N}}(\lambda_{\min}(P_i))\|e(t)\|^2$$

one obtains

$$\dot{V}(t) \leq -\mu\min_{i\in\mathbf{N}}(\lambda_{\min}(P_i))\|e(t)\|^2 + 4\sum_{i=1}^{N}\sum_{q=1}^{p_i}|H_{i,q}|\beta_{i,q}^2.$$

Therefore, $e(t)$ is uniformly ultimately bounded with ultimate bound

$$\lim_{t\to\infty}\sup\|e(t)\| \leq 2\sqrt{\frac{\sum_{i=1}^{N}\sum_{q=1}^{p_i}|H_{i,q}|\beta_{i,q}^2}{\mu\min_{i\in\mathbf{N}}(\lambda_{\min}(P_i))}}.$$

Therefore, the proof is completed. ■

Remark 4: It should be noted that according to (46) and (47), the time delay is not in the output injection loop of the observer, because the delay just affects the output measurement of the plant. As the delay is fixed, the delay value does not have any effect on the stability and convergence analysis of the observer. However, the delay leads to the convergence of the estimated state vector to a delayed state vector.

Remark 5: The ultimate bound of the estimation errors depends on the accuracy of the sensors. In this study, in (4) we have considered the bounds $\beta_{i,q}, q \in \mathbf{p}_i, i \in \mathbf{N}$, for faults and noise acceptable for practically healthy sensors. If Assumption 3 is satisfied, the ultimate bound of the estimation errors is a function of $\beta_{i,q}, q \in \mathbf{p}_i, i \in \mathbf{N}$. This implies that if healthy measurements are noiseless, the estimation error vector e asymptotically converges to zero even in the presence of fault in unhealthy sensors if Assumption 3 is verified.

Remark 6: By using the robustifying term $\nu_i(t)$, any entries of the second term in the right-hand side of (54) are nonpositive as (55) or are bounded as (56). Without using the robustifying term, the boundedness of the second term in the right-hand side of (54) should be guaranteed by the term $-\mu V(t)$, and hence in such case, the ultimate bound (48) may not be obtained.

Remark 7: From (57) we have

$$\dot{V}(t) \leq -\mu V(t) + d$$

where

$$d(t) = 4 \sum_{i=1}^{N} \sum_{q=1}^{p_i} |H_{i,q}| \beta_{i,q}^2.$$

From the comparison theorem for scalar ordinary differential equations [42], one obtains $V(t) \leq v(t)$ for all $t \geq 0$, where $v(t)$ is given by

$$v(t) = e^{-\mu t} V(0) + \frac{d}{\mu} \left(1 - e^{-\mu t}\right)$$

implying the convergence of $v(t)$ to d/μ with time constant $1/\mu$.

VI. Simulation Results

The effectiveness of the proposed distributed observer design is evaluated via numerical examples. We consider a jointly observable Lipschitz nonlinear system in the form of (1) when

$$A = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 0 & 3 \\ 0 & -2 & 0 \end{bmatrix}, \; B = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

and by defining $x = \begin{bmatrix} x^1 & x^2 & x^3 \end{bmatrix}^\top$, the Lipschitz nonlinear term is given by

$$f(x) = \begin{bmatrix} 1.2 \sin(x^1) \\ 0 \\ 0.8 \sin(x^2) \cos(x^3) \end{bmatrix}.$$

The Lipschitz constant γ is selected as $\gamma = 2$. Define

$$C_1 = [0 \;\; 0 \;\; 1], \; C_2 = [0 \;\; 0 \;\; 0], \; C_3 = [0 \;\; 1 \;\; 1]$$

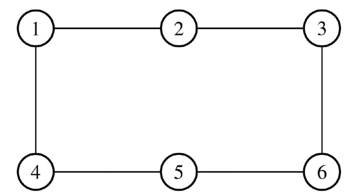


Fig. 5. Network communication graph.

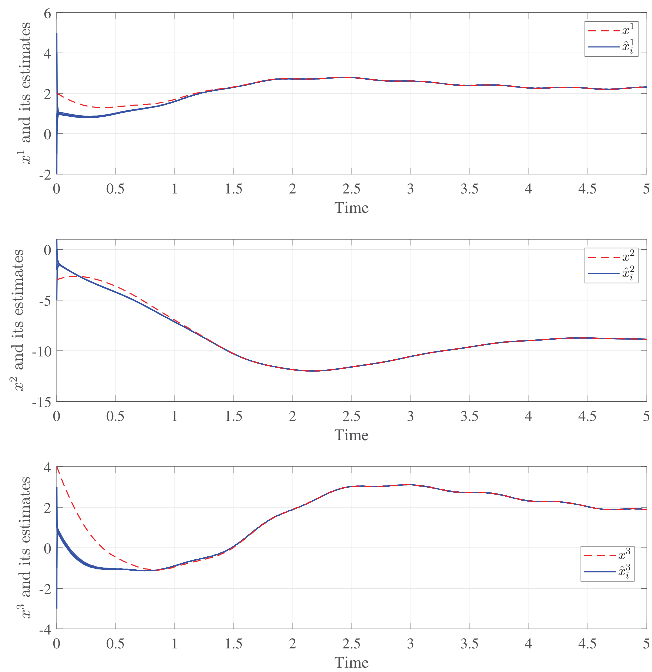


Fig. 6. Estimated state and plant state vectors for the noise-free and fault-free scenario (no sensor redundancy employed).

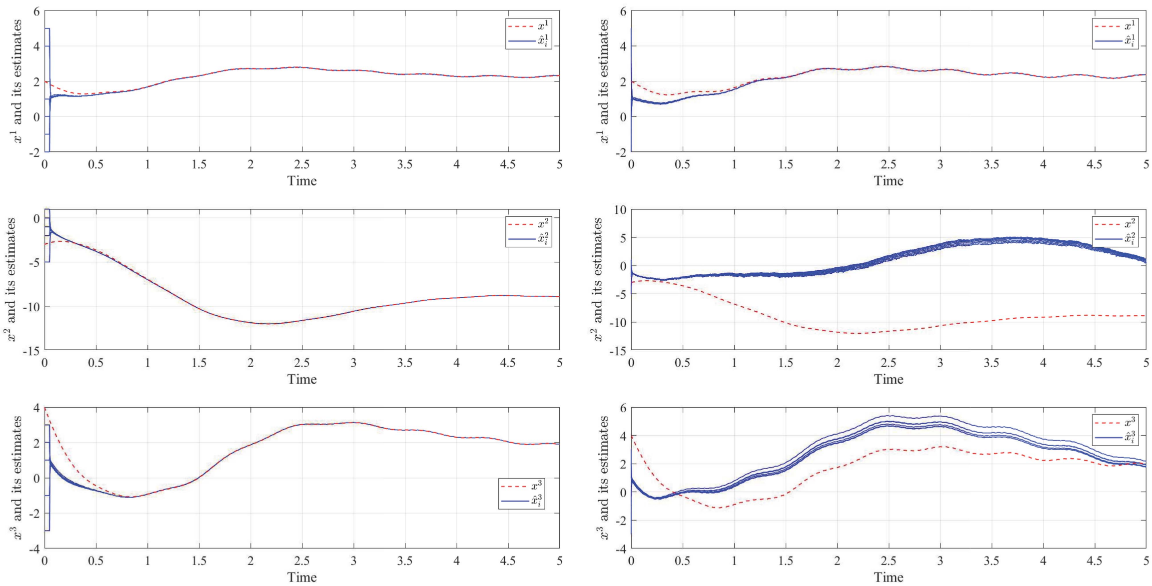
$$C_4 = [1 \;\; 0 \;\; 0], \; C_5 = [0 \;\; 0 \;\; 0], \; C_6 = [0 \;\; 1 \;\; 0]$$

and let three measurements at each node be considered, i.e., $r_i = 3$, $i \in \{1, 2, \ldots, 6\}$. Note that none of the pairs (C_i, A), $i \in \{1, 2, \ldots, 6\}$, are observable. The nodes are assumed to be connected by the unweighted undirected communication graph depicted in Fig. 5 implying that $\lambda_2(\mathcal{L}) = 1$.

Following Assumption 3, we assume that to measure each output $y_i, i \in \{1, 2, \ldots, 6\}$, two sensors are considered healthy with level $\boldsymbol{\beta_i} = 0.6$, $i \in \{1, 2, \ldots, 6\}$, whereas another sensor suffers from an additive fault with a magnitude larger than 50. Note that since $p_i = 1$, $i \in \{1, 2, \ldots, 6\}$, $\boldsymbol{\beta_i}$ and $\boldsymbol{\eta_i}$ in (47) are both scalars. A variety of sinusoidal waves, square waves, and uniformly random signals are considered to model the noise and faults. For the given matrices A and C_i, the similarity transformation matrices T_i read as

$$T_1 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \; T_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \; T_3 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$T_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \; T_5 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \; T_6 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$



(a) (b)

Fig. 7. (a) Estimated state and plant state vectors in the presence of noise and sensor faults using sensor redundancy. (b) Estimated state and plant state vectors in the presence of noise and sensor faults without sensor redundancy.

It is worth mentioning that there are no T_{io} for Node 2 and Node 5 since they have no measurements.

For the distributed observer in Theorem 2, the matrices P_i and H_i and the injection gain L_i are obtained from a numerical solution of the LMI (14), which yields

$$\begin{aligned}
&P_1 = P_3 = P_6 = \mathrm{diag}(1,\ 2.84,\ 2.84)\\
&P_2 = P_5 = I_3,\ P_4 = \mathrm{diag}(2.85,\ 1,\ 1)\\
&H_1 = -57.34,\ H_2 = H_5 = 0\\
&H_3 = -5.76,\ H_4 = -82.04,\ H_6 = -57.34\\
&L_1 = [0 \quad 0 \quad -20.17]^\top\\
&L_2 = L_5 = \mathbf{0}_{3\times 1},\\
&L_3 = [0 \quad -2.03 \quad -2.03]^\top\\
&L_4 = [-28.82 \quad 0 \quad 0]^\top\\
&L_6 = [0 \quad -20.17 \quad 0]^\top.
\end{aligned} \tag{58}$$

Moreover, we set $\boldsymbol{\eta_i} = 0.6,\ i \in \{1, 2, \ldots, 6\}$, and following (16), χ is set as $\chi = \|\Lambda + \gamma\bar{P} + \Lambda_P^\top Q^{-1}\Lambda_P\|/(2\lambda_2(\mathcal{L})) + 1$.

Before investigating the effectiveness of the estimation strategy of Theorem 2 in tolerating sensor faults, we show that it is possible to estimate the plant state by employing the estimation strategy of Theorem 1. For this preliminary noise-free and fault-free scenario, the simulation results reported in Fig. 6 show the effectiveness of the proposed distributed observer when just one cluster of sensors is used.

Next, we investigate the estimation strategy of Theorem 2. In the simulation, the fixed time delay for Algorithm 1 has been set as $\tau = 5 \times 10^{-2}$ and $\dot{\hat{x}}_i(t) = 0$ when $t < \tau$ for all $i \in \mathbf{N}$. According to the simulation results illustrated in Fig. 7(a), all the local observers are capable of estimating the state vector. Moreover, we notice from the simulation results that the effects of the severe faults with large magnitude are rejected without any abnormal behavior in the observers. However, when Assumption 3 is not satisfied, state estimation may be sensitive to sensor faults. To show this issue, we repeat the first scenario when just one cluster of sensors (selected randomly) are employed in the observers. The simulation results for this scenario are depicted in Fig. 7(b), implying that the performance of the distributed observer in state estimation has significantly deteriorated. Note that, according to (46) and (47), the discontinuous function $\mathrm{msgn}(.)$ just affects the derivative of $\hat{x}_i(t)$; therefore, the integrator between $\dot{\hat{x}}_i(t)$ and $\hat{x}_i(t)$ removes the discontinuity.

VII. Conclusion

This article proposed a robust estimation strategy for a class of nonlinear systems with distributed sensors. We designed a network of distributed observers using local sensors, under the assumption that local measurements for each observer might not be sufficient for observability. We have shown that when the observers exchange their estimated state vectors under a connected communication topology, the estimated state vector of each observer converge to the state vector of the system if the ensemble of all measurements in the network guarantees observability. In addition, when redundant sensors for each observer are available, and under suitable assumptions, the proposed estimation strategy can be modified to provide robustness against the effect of faulty sensors, without the need to employ fault detection mechanisms.

This study is regarded as a preliminary step toward the design of distributed state estimation for nonlinear systems. Many issues remain open for future research, starting from the inclusion of more general classes of nonlinear vector fields not satisfying a global Lipschitz condition. Other areas of investigation include distributed state estimation of nonlinear systems when communication among the observers is directed, distributed state estimation when the control input is unknown, and distributed state estimation in the presence of communication links failure, which are challenging problems in this area to be addressed as future work.

REFERENCES

[1] A. Mitra and S. Sundaram, "Distributed observers for LTI systems," *IEEE Trans. Autom. Control*, vol. 63, no. 11, pp. 3689–3704, Nov. 2018.

[2] L. Wang and A. S. Morse, "A distributed observer for a time-invariant linear system," *IEEE Trans. Autom. Control*, vol. 63, no. 7, pp. 2123–2130, Jul. 2018.

[3] R. Olfati-Saber, "Distributed Kalman filtering for sensor networks," in *Proc. 46th IEEE Conf. Decis. Control*, New Orleans, LA, USA, 2007, pp. 5492–5498.

[4] R. Olfati-Saber, "Kalman-consensus filter: Optimality, stability, and performance," in *Proc. 48th IEEE Conf. Decis. Control Held Jointly With 28th Chin. Control Conf.*, Shanghai, China, 2009, pp. 7036–7042.

[5] M. Kamgarpour and C. Tomlin, "Convergence properties of a decentralized Kalman filter," in *Proc. 47th IEEE Conf. Decis. Control*, Cancun, Mexico, 2008, pp. 3205–3210.

[6] B. Shen, Z. Wang, and Y. S. Hung, "Distributed $\mathcal{H}_\infty$-consensus filtering in sensor networks with multiple missing measurements: The finite-horizon case," *Automatica*, vol. 46, no. 10, pp. 1682–1688, Oct. 2010.

[7] V. Ugrinovskii, "Distributed robust estimation over randomly switching networks using $\mathcal{H}_\infty$ consensus," *Automatica*, vol. 49, no. 1, pp. 160–168, Jan. 2013.

[8] M. Farina, G. Ferrari-Trecate, and R. Scattolini, "Distributed moving horizon estimation for linear constrained systems," *IEEE Trans. Autom. Control*, vol. 55, no. 11, pp. 2462–2475, Nov. 2010.

[9] I. Matei and J. S. Baras, "Consensus-based linear distributed filtering," *Automatica*, vol. 48, no. 8, pp. 1776–1782, Aug. 2012.

[10] S. Park and N. C. Martins, "Design of distributed linear time-invariant observers for state omniscience," *IEEE Trans. Autom. Control*, vol. 62, no. 2, pp. 561–576, Feb. 2017.

[11] L. Wang, J. Liu, A. S. Morse, and B. Anderson, "A distributed observer for a discrete-time linear system," in *Proc. 58th IEEE Conf. Decis. Control*, Nice, France, 2019, pp. 367–372.

[12] A. Mitra, J. A. Richards, S. Bagchi, and S. Sundaram, "Resilient distributed state estimation with mobile agents: Overcoming Byzantine adversaries, communication losses, and intermittent measurements," *Auton. Robots*, vol. 43, pp. 743–768, Mar. 2019.

[13] A. Mitra and S. Sundaram, "Byzantine-resilient distributed observers for LTI systems," *Automatica*, vol. 108, pp. 1–12, Oct. 2019.

[14] D. Ding, Z. Wang, H. Dong, and H. Shu, "Distributed $\mathcal{H}_\infty$ state estimation with stochastic parameters and nonlinearities through sensor networks: The finite-horizon case," *Automatica*, vol. 48, no. 8, pp. 1575–1585, Aug. 2012.

[15] J. Hu, Z. Wang, J. Liang, and H. Dong, "Event-triggered distributed state estimation with randomly occurring uncertainties and nonlinearities over sensor networks: A delay-fractioning approach," *J. Franklin Inst.*, vol. 352, no. 9, pp. 3750–3763, Sep. 2015.

[16] J. Liang, Z. Wang, T. Hayat, and A. Alsaedi, "Distributed $\mathcal{H}_\infty$ state estimation for stochastic delayed 2-D systems with randomly varying nonlinearities over saturated sensor networks," *Inf. Sci.*, vol. 370–371, pp. 708–724, Nov. 2016.

[17] J. Liang, Z. Wang, and X. Liu, "Distributed state estimation for discrete-time sensor networks with randomly varying nonlinearities and missing measurements," *IEEE Trans. Neural Netw.*, vol. 22, no. 3, pp. 486–496, Mar. 2011.

[18] S. Wang, H. Fang, and X. Liu, "Distributed state estimation for stochastic non-linear systems with random delays and packet dropouts," *IET Control Theory Appl.*, vol. 9, no. 18, pp. 2657–2665, Dec. 2015.

[19] D. Wang, Z. Wang, G. Li, and W. Wang, "Distributed filtering for switched nonlinear positive systems with missing measurements over sensor networks," *IEEE Sensors J.*, vol. 16, no. 12, pp. 4940–4948, Jun. 2016.

[20] M. S. Chong, H. Sandberg, and J. P. Hespanha, "A secure state estimation algorithm for nonlinear systems under sensor attacks," in *Proc. 59th IEEE Conf. Decis. Control*, Jeju, South Korea, 2020, pp. 5743–5748.

[21] X. He, X. Zhang, W. Xue, and H. Fang, "Distributed Kalman filter for a class of nonlinear uncertain systems: An extended state method," in *Proc. 21st Int. Conf. Inf. Fusion*, Cambridge, U.K., 2018, pp. 78–83.

[22] X. He, W. Xue, X. Zhang, and H. Fang, "Distributed filtering for uncertain systems under switching sensor networks and quantized communications," *Automatica*, vol. 114, pp. 1–13, Apr. 2020.

[23] R. Rajamani, "Observers for Lipschitz nonlinear systems," *IEEE Trans. Autom. Control*, vol. 43, no. 3, pp. 397–401, Mar. 1998.

[24] G. Phanomchoeng and R. Rajamani, "Observer design for Lipschitz nonlinear systems using Riccati equations," in *Proc. Amer. Control Conf.*, Baltimore, MD, USA, 2010, pp. 6060–6065.

[25] G. Yang, H. Rezaee, and T. Parisini, "Distributed state estimation for a class of jointly observable nonlinear systems," in *Proc. 21st IFAC World Congr.*, Berlin, Germany, 2020, pp. 5045–5050.

[26] D. Angeli, B. Ingalls, E. D. Sontag, and Y. Wang, "Separation principles for input-output and integral-input-to-state stability," *SIAM J. Control Optim.*, vol. 43, no. 1, pp. 256–276, 2004.

[27] W. Ren, R. W. Beard, and E. M. Atkins, "Information consensus in multivehicle cooperative control," *IEEE Control Syst. Mag.*, vol. 27, no. 2, pp. 71–82, Apr. 2007.

[28] T. Kim, H. Shim, and D. D. Cho, "Distributed Luenberger observer design," in *Proc. 55th IEEE Conf. Decis. Control*, Las Vegas, NV, USA, 2016, pp. 6928–6933.

[29] W. M. Wonham, *Linear Multivariable Control a Geometric Approach*, 3rd ed. New York, NY, USA: Springer, 1985.

[30] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1520–1533, Sep. 2004.

[31] S. P. Boyd, L. E. Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. Philadelphia, PA, USA: SIAM, 1994.

[32] W. Han, H. L. Trentelman, Z. Wang, and Y. Shen, "A simple approach to distributed observer design for linear systems," *IEEE Trans. Autom. Control*, vol. 64, no. 1, pp. 329–336, Jan. 2019.

[33] F. R. S. Sevilla, I. M. Jaimoukha, B. Chaudhuri, and P. Korba, "A semidefinite relaxation procedure for fault-tolerant observer design," *IEEE Trans. Autom. Control*, vol. 60, no. 12, pp. 3332–3337, Dec. 2015.

[34] C. Zhang, I. M. Jaimoukha, and F. R. S. Sevilla, "Fault-tolerant observer design with a tolerance measure for systems with sensor failures," in *Proc. Amer. Control Conf.*, Boston, MA, USA, 2016, pp. 7523–7528.

[35] S. S. Nasrolahi and F. Abdollahi, "Sensor fault detection and recovery in satellite attitude control," *Acta Astronautica*, vol. 145, pp. 275–283, Apr. 2018.

[36] K. Rothenhagen and F. W. Fuchs, "Current sensor fault detection, isolation, and reconfiguration for doubly fed induction generators," *IEEE Trans. Ind. Electron.*, vol. 56, no. 10, pp. 4239–4245, Oct. 2009.

[37] Y. Peng, W. Qiao, L. Qu, and J. Wang, "Sensor fault detection and isolation for a wireless sensor network-based remote wind turbine condition monitoring system," *IEEE Trans. Ind. Appl.*, vol. 54, no. 2, pp. 1072–1079, Mar./Apr. 2018.

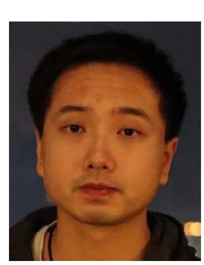
[38] H. Behzad, A. Casavola, F. Tedesco, and M. A. Sadrnia, "Fault-tolerant sensor reconciliation schemes based on unknown input observers," *Int. J. Control*, vol. 93, no. 3, pp. 669–679, 2020.

[39] J.-J. E. Slotine, J. K. Hedrick, and E. A. Misawa, "On sliding observers for nonlinear systems," *J. Dyn. Syst., Meas., Control*, vol. 109, no. 3, pp. 245–252, Sep. 1987.

[40] T. Ito, "A Filippov solution of a system of differential equations with discontinuous right-hand sides," *Econ. Lett.*, vol. 4, no. 4, pp. 349–354, 1979.

[41] A. F. Filippov, *Differential Equations With Discontinuous Righthand Sides*. Dordrecht, The Netherlands: Springer, 1988.

[42] I. Karafyllis, "Lyapunov theorems for systems described by retarded functional differential equations," *Nonlinear Anal.: Theory, Methods Appl.*, vol. 64, no. 3, pp. 590–617, 2006.

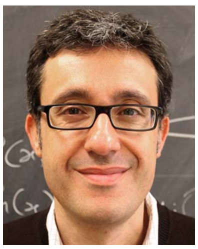


Guitao Yang received the B.Eng. degree in electrical and electronic engineering from the University of Manchester, Manchester, U.K., in 2016, and the M.Sc. degree in control systems, in 2017, from Imperial College London, London, U.K., where he is currently working toward the Ph.D. degree with the Department of Electrical and Electronic Engineering.

His research interests include distributed state estimation, fault-tolerant observers, and geometric-based state estimation.

Hamed Rezaee (Member, IEEE) received the B.Sc., M.Sc., and Ph.D. degrees in control engineering from the Department of Electrical Engineering, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran, in 2009, 2011, and 2016, respectively.

He is currently a Research Associate with the Department of Electrical and Electronic Engineering, Imperial College London, London, U.K. His research interests include monitoring and resilient control in cyber-physical systems, multiagent systems, connected vehicles, and distributed state estimation.



Andrea Serrani (Member, IEEE) received the laurea (B.Eng.) degree in electrical engineering, summa cum laude and the Ph.D. degree in artificial intelligence systems from the University of Ancona, Italy, in 1993 and 1997, and the M.S. and D.Sc. degrees in systems science and mathematics from Washington University, St. Louis, MO, USA, in 1996 and 2000, respectively.

From 1994 to 1999, he was a Fulbright Fellow with Washington University. Since 2002, he has been with the Department of Electrical and Computer Engineering, Ohio State University, Columbus, OH, USA, where he is currently a Professor and Associate Chair. He has held visiting positions with the University of Bologna, Bologna, Italy, and University of Padua, Padua, Italy, and multiple summer faculty fellowships with AFRL. His work has been supported by AFRL, NSF, Ford Motor Co., and NASA, among others. He has authored or coauthored more than 150 articles in journals, proceedings of international conferences, and book chapters, and is the coauthor of the book *Robust Autonomous Guidance: An Internal Model Approach* (Springer-Verlag). His research interests include methodological aspects of nonlinear, adaptive and geometric control theory with applications in aerospace and marine systems, fluidic systems, robotics and automotive engineering.

Prof. Serrani is the Editor-in-Chief for the IEEE TRANSACTIONS ON CONTROL SYSTEMS TECHNOLOGY and was an Associate Editor for the IEEE TRANSACTIONS ON CONTROL SYSTEMS TECHNOLOGY, from 2010 to 2016, *Automatica*, from 2008 to 2014, and the *International Journal of Robust and Nonlinear Control*, from 2006 to 2014. He serves on the Conference Editorial Boards of IEEE Control Systems Society (CSS) and International Federation of Automatic Control, served as Program Chair for the 2019 American Control Conference, and as General Co-chair for the 2022 IEEE Conference on Decision and Control. He was a Distinguished Lecturer of the IEEE CSS.

Thomas Parisini (Fellow, IEEE) received the Ph.D. degree in electronic engineering and computer science from the University of Genoa, Genoa, Italy, in 1993.

He was with Politecnico di Milano, Milano, Italy, and since 2010, he holds the Chair of Industrial Control and is Director of Research with Imperial College London, U.K. He is the Deputy Director of the KIOS Research and Innovation Centre of Excellence, University of Cyprus, Nicosia, Cyprus. Since 2001, he is also Danieli Endowed Chair of Automation Engineering with the University of Trieste, Trieste, Italy, where he was the Deputy Rector, from 2009 to 2012. In 2018, he received an Honorary Doctorate from the University of Aalborg, Aalborg, Denmark. In 2020, he has been appointed as Deputy Chair of the Employment and Education Task Force with the B20–Italy. He has authored or coauthored more than 350 research papers in archival journals, book chapters, and international conference proceedings.

Prof. Parisini is a Fellow of the IFAC. He was the co-recipient of the IFAC Best Application Paper Prize of the Journal of Process Control, Elsevier, from 2011 to 2013, and the 2004 Outstanding Paper Award of the IEEE TRANSACTIONS ON NEURAL NETWORKS. He was the recipient of the 2007 IEEE Distinguished Member Award. Since 2017, he is Editor for Control Applications of Automatica and since 2018 he is the Editor-in-Chief of the *European Journal of Control*. During 2009 to 2016, he was the Editor-in-Chief for the IEEE TRANSACTIONS ON CONTROL SYSTEMS TECHNOLOGY. In 2016, he was awarded as Principal Investigator at Imperial of the H2020 European Union flagship Teaming Project KIOS Research and Innovation Centre of Excellence led by University of Cyprus. He serves as the President of the IEEE Control Systems Society, from 2021 to 2022, and has served as the Vice-President for Publications Activities. Among other activities, he was the Program Chair of the 2008 IEEE Conference on Decision and Control and General Co-Chair of the 2013 IEEE Conference on Decision and Control.