

Distributed Fault Diagnosis for Continuous-Time Nonlinear Systems: the Input-Output case

Francesca Boem^a Riccardo M. G. Ferrari^b Thomas Parisini^c
Marios M. Polycarpou^d

^a *University of Trieste, Italy (e-mail: francesca.boem@phd.units.it).*

^b *Danieli Automation S.p.A., Buttrio, Italy (e-mail: r.ferrari@dca.it).*

^c *Imperial College London, UK & University of Trieste, Italy (e-mail: t.parisini@gmail.com).*

^d *KIOS Research Center, ECE Department, University of Cyprus, Nicosia 1678, Cyprus (e-mail: mpolycar@ucy.ac.cy)*

Abstract

In this paper, new results on distributed fault diagnosis of continuous-time nonlinear systems with partial state measurements are proposed. Following an overlapping decomposition framework, the dynamics of a nonlinear uncertain large-scale dynamical system are described as the interconnections of several subsystems. Each subsystem is monitored by a corresponding Local Fault Diagnoser: a set of local estimators, based on the nominal local dynamic model and on an adaptive approximation of the interconnection and of the fault function, permit to derive a local fault decision. A consensus-based protocol is used in order to improve the detectability and the isolability of faults affecting variables shared among different subsystems because of the overlapping decomposition. A sufficient condition assuring the convergence of the estimation errors is derived. Time-varying threshold functions guaranteeing no false-positive alarms and theoretical results containing detectability and isolability conditions are presented.

Key words: Distributed Fault Detection and Isolation, Non-linear systems, Large-scale systems.

1 Introduction

The growing scientific interest for distributed systems and networks is testified by the considerable research accomplished, which is cited in surveys such as [1], [2] and [3]. The scientific literature in this field can be divided into two main groups. The first one deals with the modeling and the analysis of existing distributed systems and covers a variety of different scientific areas, such as psychology, sociology and biology. The second group, which is of greater interest to the control community, includes works directed at designing distributed systems, such as the notable contributions of [4], [5] and [6], among many others. The goal is to find distributed solutions to problems that are difficult or impossible to solve in a centralized way, because of structural, efficiency, computational, and robustness issues. Practical engineering examples of large-scale and/or distributed systems

are abundant, including large-scale communication networks, water distribution systems, traffic networks, energy, pulp-and-paper, steel-making plants, multi-vehicle formations, and so on. "Decentralized control" methods suited to these systems were proposed since at least the 1970s, as described in the survey work of [7]. Although many enhancements have been achieved in the design and analysis of decentralized and, later, distributed control and estimation schemes, the design of fault diagnosis schemes specifically for distributed and large-scale systems is still at an early stage. Research activity in this field led to effective distributed algorithms suited to discrete event systems (see, among many others, [8] and [9]). A notable contribution in the field of decentralized hybrid systems fault diagnosis is the work of [10], while an interesting scheme for the nonlinear fault detection of spacecraft formations was presented in [11]. Furthermore, analysis of fault scenarios and effects in distributed

systems were addressed in [12] and in [13]. However, as far as distributed discrete-time or continuous-time systems are concerned, only qualitative fault diagnosis schemes were attempted very recently [14], [15], or quantitative methods that were formulated for linear systems only [16], [17], with [18], [19] being some of the few contributions on decentralized/distributed fault detection for large-scale nonlinear systems.

In this paper, we propose a fault detection and isolation scheme for a class of non-linear uncertain continuous-time systems, where, furthermore, the system states are only partially measurable (see a preliminary version in [20]). The diagnosis architecture was proposed for the full-state case in [21] and [22] for continuous-time and discrete-time systems, respectively. Though several papers dealing with centralized fault diagnosis schemes for input-output systems have appeared in the literature ([23–27]), the contributions addressing distributed schemes for input-output nonlinear systems are few ([28], dealing with discrete-time systems, and [29]). More specifically, the main contributions of this paper are the formulation of a distributed architecture specifically for input-output continuous-time systems, with a sufficient condition guaranteeing the convergence of the estimators, and the derivation of rigorous analytical conditions for detectability and isolability. The paper is organized as follows. Section 2 formulates the problem under investigation. A distributed fault detection architecture is presented in Section 3, where a sufficient condition assuring the convergence of the estimation errors is derived, followed by the derivation of detectability conditions. After that, in Section 4, the isolation problem is considered and theoretical results regarding fault isolability are presented. Finally, Section 5 reports some concluding remarks.

2 Problem formulation

We consider a multi-input multi-output uncertain nonlinear system, called the *monolithic* system:

$$\mathcal{S} : \begin{cases} \dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{f}(\mathbf{x}, \mathbf{u}) + \boldsymbol{\eta}_x(\mathbf{x}, \mathbf{u}, t) + \beta(t - T_0)\boldsymbol{\phi}(\mathbf{x}, \mathbf{u}) \\ \mathbf{y} = \mathbf{C}\mathbf{x} + \boldsymbol{\eta}_y(\mathbf{x}, \mathbf{u}, t), \end{cases} \quad (1)$$

where ¹ $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{u} \in \mathbb{R}^m$ and $\mathbf{y} \in \mathbb{R}^p$ denote the state, the control input and the measured output vectors respectively, the matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ and the vector field $\mathbf{f} : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^n$ describe the nominal healthy dynamics, $\mathbf{C} \in \mathbb{R}^{p \times n}$ is the nominal output matrix, $\boldsymbol{\eta}_x$ and $\boldsymbol{\eta}_y$ are the uncertainties in the state and in the output equations. The term $\beta(t - T_0)\boldsymbol{\phi}(\mathbf{x}, \mathbf{u})$ represents the fault function dynamics: $\boldsymbol{\phi}(\mathbf{x}, \mathbf{u})$ denotes the functional structure and $\beta(t - T_0)$ characterizes the time profile of

the fault:

$$\beta(t - T_0) \triangleq \begin{cases} 0 & \text{if } t < T_0 \\ 1 - e^{-\alpha(t - T_0)} & \text{if } t \geq T_0 \end{cases} \quad (2)$$

where T_0 is the unknown fault occurrence time and $\alpha > 0$ represents the unknown fault-evolution rate, modeling either *incipient* faults or *abrupt* faults, as $\alpha \rightarrow \infty$. The following assumptions are needed.

Assumption 1 *The state variables \mathbf{x} and control variables \mathbf{u} are bounded before and after the occurrence of a fault: $\exists \mathcal{R}$, compact region of $\mathbb{R}^n \times \mathbb{R}^m$: $(\mathbf{x}(t), \mathbf{u}(t)) \in \mathcal{R}$, $\forall t \geq 0$.*

Assumption 2 *The fault-evolution rate parameter α is unknown, but lower bounded by a known constant $\bar{\alpha}$.*

Assumption 3 *The measurement uncertainty term $\boldsymbol{\eta}_y$ is an unstructured and unknown nonlinear function, bounded by a positive, known and bounded function $\bar{\boldsymbol{\eta}}_y$:*

$$\left| \boldsymbol{\eta}_y^{(k)}(\mathbf{x}, \mathbf{u}, t) \right| \leq \bar{\boldsymbol{\eta}}_y^{(k)}(\mathbf{x}, \mathbf{u}, t),$$

for all $k = 1, \dots, p$, $(\mathbf{x}, \mathbf{u}) \in \mathcal{R}$ and for all $t \geq 0$.

In this paper, the methodology presented in [28] for discrete-time systems is adapted to the continuous-time context, based on the decomposition of the monolithic system \mathcal{S} into N subsystems \mathcal{S}_I , $I = 1, \dots, N$. It is important to note that in the present paper, for the sake of generality, we consider the decomposition of the *structural graph* ([30]), whose nodes are the state and the input variables of the system, instead of a decomposition made only with respect to the measurable variables. The decomposition allows the overlapping of certain states, that is, certain state variables may belong to more than one subsystem. We refer to them as *shared* variables. After the decomposition, the I -th subsystem \mathcal{S}_I dynamics can be described by:

$$\begin{cases} \dot{x}_I = A_I x_I + f_I(x_I, u_I) + g_I(C_I x_I, u_I, z_I) \\ \quad + \beta(t - T_0)\boldsymbol{\phi}_I(C_I x_I, z_I, u_I) \\ y_I = C_I x_I + \eta_{y,I}(x_I, u_I, t), \end{cases} \quad (3)$$

where $x_I \in \mathbb{R}^{n_I}$, $u_I \in \mathbb{R}^{m_I}$ and $y_I \in \mathbb{R}^{p_I}$ are the local state, the local control input, and the local measured output vectors respectively, and $z_I \in \mathbb{R}^{q_I}$ is the vector of the interconnection variables². The matrix $A_I \in \mathbb{R}^{n_I \times n_I}$ and the vector field $f_I : \mathbb{R}^{n_I} \times \mathbb{R}^{m_I} \mapsto \mathbb{R}^{n_I}$ represent the local nominal healthy dynamics, $C_I \in \mathbb{R}^{p_I \times n_I}$ is the nominal local output matrix, $g_I : \mathbb{R}^{p_I} \times \mathbb{R}^{m_I} \times \mathbb{R}^{q_I} \mapsto \mathbb{R}^{n_I}$

¹ The use of boldface letters indicates that a given variable is related to the system (1).

² The interconnection variables z_I are the states of the neighboring subsystems nodes in the structure graph having a connection with the subsystem I .

is the interconnection function that incorporates the effects of the corresponding local state modeling uncertainty term $\eta_{x,I}$. The term $\eta_{y,I}$ is the uncertainty function in the local output equation that takes into account the measurement error, while $\phi_I : \mathbb{R}^{p_I} \times \mathbb{R}^{m_I} \times \mathbb{R}^{q_I} \mapsto \mathbb{R}^{n_I}$ is the local fault function. The following assumptions are in place.

Assumption 4 *The decomposition of the monolithic system (1) is such that z_I is made of measurable variables only.*

This assumption is needed in order to allow the learning of the interconnection function and of the fault function.

Assumption 5 *The structural graph and the decomposition are the same before and after the fault event.*

In this way we suppose that the possible fault event does not cause a change to the system structure by adding new dependencies between variables belonging to different subsystems. However, it is possible for a fault event to remove some of the interconnections, which can be formally represented by setting some g_I function to zero.

Assumption 6 *(A_I, C_I) is an observable pair, $\forall I = 1, \dots, N$.*

Assumption 7 *The interconnection function g_I is an unstructured and uncertain nonlinear function, bounded by a known and bounded function, i.e., $|g_I^{(k)}(C_I x_I, z_I, u_I)| \leq \bar{g}_I^{(k)}(C_I x_I, z_I, u_I)$, for all $I = 1, \dots, N$, $k = 1, \dots, n_I$ and for all $(x, u) \in \mathcal{R}$.*

3 Distributed Detection Architecture

In [22], a *Distributed Fault Detection and Isolation* (DFDI) methodology is proposed, consisting of N agents called *Local Fault Diagnosers* (LFDs) \mathcal{L}_I , $I \in \{1 \dots N\}$. For detection purposes, each LFD is equipped with a non-linear adaptive estimator, called *Fault Detection Approximation Estimator* (FDAE), based on the nominal model, that computes the estimate of the local state x_I and of the local output y_I . The difference between the estimated output \hat{y}_I and the measurements y_I is the output estimation error $\epsilon_{y,I} \triangleq y_I - \hat{y}_I$, which plays the role of a residual and will be compared, component by component, to a suitable detection threshold signal $\bar{\epsilon}_{y,I}$. The following

$$\left| \epsilon_{y,I}^{(k)}(t) \right| \leq \bar{\epsilon}_{y,I}^{(k)}(t), \quad \forall k = 1, \dots, p_I \quad (4)$$

is a necessary (but generally not sufficient) condition for the fault-free hypothesis \mathcal{H}_I : “The system \mathcal{S}_I is healthy”. If the condition is violated at some time instant t , then the hypothesis \mathcal{H}_I is falsified.

Definition 3.1 *The local fault detection time is defined as $T_{d,I} = \min \left\{ t : \exists k, k \in 1, \dots, p_I, \left| \epsilon_{y,I}^{(k)}(t) \right| > \bar{\epsilon}_{y,I}^{(k)}(t) \right\}$.*

The local FDAE estimation, in the case of non-shared state variables, can be computed as:

$$\begin{aligned} \dot{\hat{x}}_I &= A_I \hat{x}_I + f_I(\hat{x}_I, u_I) + \hat{g}_I(y_I, u_I, v_I, \hat{\vartheta}_I) \\ &\quad + L_I(y_I - \hat{y}_I) \\ \hat{y}_I &= C_I \hat{x}_I, \end{aligned} \quad (5)$$

where \hat{g}_I is the output of an adaptive approximator designed to learn the unknown interconnection function g_I and $\hat{\vartheta}_I \in \hat{\Theta}_I$ denotes its adjustable parameters vector. The learning law will be derived in the following. Due to the uncertain output measurements, each LFD receives from its neighbors the vector $v_I = z_I + \varsigma_I$, where ς_I is made with the components of $\eta_{y,J}$ that affect the relevant components of the neighboring subsystems measurements y_J . In the case of variables $x^{(s)}$ shared among more than one LFD, we take advantage of the redundancy obtained by means of the overlap, by proposing a deterministic consensus protocol defined on a generic communication graph $\mathcal{G}_s \triangleq (\mathcal{O}_s, \mathcal{E}_s)$, whose nodes are the LFDs in the overlap set \mathcal{O}_s of $x^{(s)}$ (see [22]):

$$\begin{aligned} \dot{\hat{x}}_I^{(s_I)} &= \sum_{J \in \mathcal{O}_s} W_s^{(I,J)} \left[A_J^{(s_J)} \hat{x}_J + f_J^{(s_J)}(\hat{x}_J, u_J) \right. \\ &\quad \left. + \hat{g}_J^{(s_J)}(y_J, u_J, v_J, \hat{\vartheta}_J) + L_J^{(s_J)}(y_J - \hat{y}_J) \right] \end{aligned} \quad (6)$$

where the terms $W_s^{(I,J)}$ are the components of a doubly stochastic weighted adjacency matrix, as for instance the Metropolis matrix ([31]):

$$W_s^{(I,J)} \triangleq \begin{cases} 0 & (I, J) \notin \mathcal{E}_s \\ \frac{1}{1 + \max\{d_s^{(I)}, d_s^{(J)}\}} & (I, J) \in \mathcal{E}_s, I \neq J \\ 1 - \sum_{K \neq I} W_s^{(I,K)} & I = J \end{cases} \quad (7)$$

with $d_s^{(I)}$ being the degree of the I -th node in the communication graph \mathcal{G}_s . Other choices are possible, but in all cases the weights $W_s^{(I,J)}$ can be seen as a level of how much the I -th subsystem feels confident about the information received from subsystem J . It is important to note that, in order to implement (6), the I -th LFD does not need the information about the expressions of $A_J^{(s_J)}$, $f_J^{(s_J)}$, $\hat{g}_J^{(s_J)}$ and of $L_J^{(s_J)}$: it is sufficient that each LFD computes locally the term $A_J^{(s_J)} \hat{x}_J + f_J^{(s_J)} + \hat{g}_J^{(s_J)} + L_J^{(s_J)}(y_J - \hat{y}_J)$ and communicates it to other LFDs according to the communication graph \mathcal{G}_s .

We now analyze the dynamics of the FDAE estimation error before the occurrence of a fault. In the non-shared

case, the i -th state estimation error component is:

$$\begin{aligned}\dot{\epsilon}_{x,I}^{(i)} &= A_I^{(i)} x_I + f_I^{(i)}(x_I, u_I) + g_I^{(i)}(C_I x_I, u_I, z_I) \\ &\quad - A_I^{(i)} \hat{x}_I - f_I^{(i)}(\hat{x}_I, u_I) - \hat{g}_I^{(i)}(y_I, u_I, v_I, \hat{v}_I) \\ &\quad - L_I^{(i)}(y_I - \hat{y}_I) \\ &= A_{0,I}^{(i)} \epsilon_{x,I} + \Delta f_I^{(i)} + \Delta g_I^{(i)} - L_I^{(i)} \eta_{y,I},\end{aligned}\tag{8}$$

where $A_{0,I} \triangleq A_I - L_I C_I$ is a stable matrix, $\Delta f_I^{(i)} \triangleq f_I^{(i)}(x_I, u_I) - f_I^{(i)}(\hat{x}_I, u_I)$ and $\Delta g_I^{(i)} \triangleq g_I^{(i)}(C_I x_I, u_I, z_I) - \hat{g}_I^{(i)}(y_I, u_I, v_I, \hat{v}_I)$. We denote with $A^{(i)}$ the i -th row of the matrix A .

In the case of shared variables, the dynamics of the LFD state estimation error component prior to the occurrence of a fault can be written as:

$$\begin{aligned}\dot{\epsilon}_{x,I}^{(s_I)} &= \dot{x}_I^{(s_I)} - \dot{\hat{x}}_I^{(s_I)} = A_I^{(s_I)} x_I + f_I^{(s_I)}(x_I, u_I) \\ &\quad + g_I^{(s_I)}(C_I x_I, u_I, z_I) - \sum_{J \in \mathcal{O}_s} W_s^{(I,J)} \left[A_J^{(s_J)} \hat{x}_J \right. \\ &\quad \left. + f_J^{(s_J)}(\hat{x}_J, u_J) + \hat{g}_J^{(s_J)}(y_J, u_J, v_J, \hat{v}_J) \right. \\ &\quad \left. + L_J^{(s_J)}(y_J - \hat{y}_J) \right].\end{aligned}$$

By assumption it holds $\sum_{J \in \mathcal{O}_s} W_s^{(I,J)} = 1$ and, thanks to the way the model decomposition was obtained, the state estimation error component can be rewritten as:

$$\begin{aligned}\dot{\epsilon}_{x,I}^{(s_I)} &= \sum_{J \in \mathcal{O}_s} W_s^{(I,J)} \left[A_{0,J}^{(s_J)} \epsilon_{x,J} + f_J^{(s_J)}(x_J, u_J) \right. \\ &\quad \left. - f_J^{(s_J)}(\hat{x}_J, u_J) + g_J^{(s_J)}(C_J x_J, u_J, z_J) \right. \\ &\quad \left. - \hat{g}_J^{(s_J)}(y_J, u_J, v_J, \hat{v}_J) - L_J^{(s_J)} \eta_{y,J} \right] \\ &= \sum_{J \in \mathcal{O}_s} W_s^{(I,J)} \left[A_{0,J}^{(s_J)} \epsilon_{x,J} + \Delta f_J^{(s_J)} + \Delta g_J^{(s_J)} \right. \\ &\quad \left. - L_J^{(s_J)} \eta_{y,J} \right].\end{aligned}$$

We now introduce a general formulation of the state error equation for analysis purpose. To this end we define the extended state estimation error vector $\epsilon_{x,E} \in \mathbb{R}^{n_E \times 1}$, with $n_E = \sum_{J=1}^N n_J$, that is a column vector collecting the state estimation error vectors of the N sub-systems: $\epsilon_{x,E} \triangleq \text{col}(\epsilon_{x,J} : J = 1, \dots, N)$. The dynamics of $\epsilon_{x,E}$ are:

$$\dot{\epsilon}_{x,E} = W [A_{0,E} \epsilon_{x,E} + \Delta f_E + \Delta g_E - L_E \eta_{y,E}] \tag{9}$$

where W is a $N \times N$ block matrix

$$W \triangleq \begin{bmatrix} W_{1,1} & \dots & W_{1,N} \\ \dots & \dots & \dots \\ W_{N,1} & \dots & W_{N,N} \end{bmatrix},$$

such that each block $W_{I,J}$, with $J = 1, \dots, N$ and $I = 1, \dots, N$ collects the consensus weights of the subsystem I with regard to the subsystem J . The diagonal blocks $W_{I,I}$ are square diagonal matrices in $\mathbb{R}^{n_I \times n_I}$, whose s_I -th diagonal element, with $s_I = 1, \dots, n_I$, is equal to the weight $W_s^{(I,I)}$ defined in Eq. (7) if $x_I^{(s_I)}$ is a shared variable, and is equal to 1 otherwise. The matrices $W_{I,J} \in \mathbb{R}^{n_I \times n_J}$, with $J \neq I$, have non-null elements only in positions (s_I, s_J) corresponding to shared variables x_s , and here they take the value of the consensus weight $W_s^{(I,J)}$. This results in W being a symmetrical, sparse and doubly-stochastic $n_E \times n_E$ matrix. $A_{0,E}$ is a $N \times N$ diagonal block matrix:

$$A_{0,E} \triangleq \begin{bmatrix} A_{0,1} & 0 & 0 & 0 \\ 0 & A_{0,2} & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & A_{0,N} \end{bmatrix},$$

where the generic block is $A_{0,J} = A_J - L_J C_J \in \mathbb{R}^{n_J \times n_J}$, for $J = 1, \dots, N$, resulting in $A_{0,E}$ being a sparse $n_E \times n_E$ matrix. $\Delta f_E(t)$ is a $n_E \times 1$ matrix, collecting the values $\Delta f_J^{(s_J)}(t)$, for each $s_J = 1, \dots, n_J$ and for every $J = 1, \dots, N$. $\Delta g_E(t)$ is defined in an analogous way as $\Delta f_E(t)$. Furthermore, $L_E \triangleq \text{blkdiag}(L_J : J = 1, \dots, N)$ is a $N \times N$ diagonal block matrix with dimension $n_E \times p_E$, where $p_E \triangleq \sum_{J=1}^N p_J$, while $\eta_{y,E}(t)$ is a $p_E \times 1$ column vector collecting the uncertainty terms of the N subsystems: $\eta_{y,E} \triangleq \text{col}(\eta_{y,I} : J = 1, \dots, N)$.

In order to guarantee the convergence of the state estimation error, the matrix $W A_{0,E}$ has to be a stable matrix. We derived a sufficient condition assuring that all the eigenvalues of the matrix are in the negative semi-plane.

Proposition 3.1 *If $A_{0,E}$ is a diagonalizable matrix and if W is made so that the elements on the diagonal $W_{i,i} > 0.5$, then $W A_{0,E}$ is a stable matrix.*

Proof 3.1 *Since $A_{0,E}$ is a diagonal block matrix where the single blocks $A_{0,I}$, $I = 1, \dots, N$ are stable matrices, it is a stable matrix in turn. If it is a diagonal or a diagonalized matrix, the elements on the diagonal are the negative eigenvalues $-\lambda_i$. Using Gerschgorin circles ([32]) on the columns, it is possible to demonstrate that all the eigenvalues of $W A_{0,E}$ are trapped in the collection of circles centered at $-W_{i,i} \lambda_i$, with radii $\sum_{k \neq i} W_{k,i} \lambda_i = (1 - W_{i,i}) \lambda_i$, where $W_{k,i}$ is the element of W corresponding to the k -th row and i -th column, with $k, i = 1, \dots, n_E$. The condition $-W_{i,i} \lambda_i + (1 - W_{i,i}) \lambda_i < 0$ assures that all the eigenvalues of $W A_{0,E}$ have negative real part. This condition is satisfied if $W_{i,i} > 0.5$ for all $i = 1, \dots, n_E$.*

It is worth noting that assuming $A_{0,E}$ to be a diagonalizable matrix is not a restrictive assumption since we can choose $L_I, \forall I$ so that this assumption is guaranteed.

After the convergence of the state estimation error has been proved, the solution of the differential equation (9) can be written as:

$$\epsilon_{x,E}(t) = \int_0^t e^{WA_{0,E}(t-\tau)} [W\Delta f_E(\tau) + W\Delta g_E(\tau) - WL_E\eta_{y,E}(\tau)d\tau] + e^{WA_{0,E}(t)}\epsilon_{x,E}(0) \quad (10)$$

The extended output estimation error is then defined as:

$$\epsilon_{y,E} \triangleq C_E\epsilon_{x,E} + \eta_{y,E} \quad (11)$$

where $C_E \triangleq \text{blkdiag}(C_J : J = 1, \dots, N)$ is a $N \times N$ diagonal block matrix, with dimension $p_E \times n_E$. From (9), (11) and the definition of C_E , the following learning law for the adjustable parameter vector $\hat{\vartheta}_I$ of the adaptive approximator $\hat{g}_I, I \in 1, \dots, N$ can be derived:

$$\begin{aligned} \dot{\hat{\vartheta}}_I &= P_{\hat{\Theta}_I} [\Gamma_I H_I^T W_{I,I}^T C_I^T \epsilon_{y,I}] \quad (12) \\ H_I^T &= \partial \hat{g}_I / \partial \hat{\vartheta}_I \end{aligned}$$

where $P_{\hat{\Theta}_I}$ is a projection operator restricting $\hat{\vartheta}_I$ within $\hat{\Theta}_I$ ([33]), Γ_I is a symmetric and positive definite learning rate matrix (see for details [21]). In the general form, the component-wise output estimation error can be bounded by the following threshold, that can be computed in a distributed way:

$$\begin{aligned} \left| \epsilon_{y,E}^{(k)}(t) \right| &\leq \left| C_E^{(k)} \epsilon_{x,E}(t) \right| + \left| \eta_{y,E}^{(k)}(t) \right| \\ &\leq \left| C_E^{(k)} \left\{ \int_0^t e^{WA_{0,E}(t-\tau)} [W\Delta f_E(\tau) + W\Delta g_E(\tau) - WL_E\eta_{y,E}(\tau)d\tau] + e^{WA_{0,E}t}\epsilon_{x,E}(0) \right\} \right| + \left| \bar{\eta}_{y,E}^{(k)}(t) \right| \\ &\leq \left| C_E^{(k)} \right| \left\{ \int_0^t \left\| e^{WA_{0,E}(t-\tau)} \right\| W \left[\bar{\Delta} f_E(\tau) + \bar{\Delta} g_E(\tau) + |L_E| \bar{\eta}_{y,E}(\tau)d\tau \right] + \left\| e^{WA_{0,E}t} \right\| \bar{\epsilon}_{x,E}(0) \right\} + \bar{\eta}_{y,E}^{(k)}(t) \\ &\triangleq \bar{\epsilon}_{y,E}^{(k)}(t) \quad (13) \end{aligned}$$

where we denote with $|A|$ the element by element absolute value of the matrix A . Furthermore,

$$\begin{aligned} \bar{\Delta} f_E^{(s)}(t) &= \max_{x^{(s)} \in R^{x^{(s)}}} \left\{ \left| \Delta f_E^{(s)}(t) \right| \right\}, \\ \bar{\epsilon}_{x,E}^{(s)}(0) &= \max_{x^{(s)} \in R^{x^{(s)}}} \left\{ \left| x^{(s)} - \hat{x}^{(s)}(0) \right| \right\}, \end{aligned}$$

for every $s = 1, \dots, n_E$. With regards to $\bar{\Delta} g_E$, some considerations are expressed in [28] and [22]: Δg_I can be upper bounded by $\bar{\Delta} g_I \triangleq \|H_I\| \kappa_I(\hat{\vartheta}_I) + \bar{\nu}_I + \max_{\eta_{y_I}} \max_{z_I} |\Delta \hat{g}_I|$, where κ_I is such that $\kappa_I(\hat{\vartheta}_I) \geq \left\| \tilde{\vartheta}_I \right\|$. In fact, by defining the parameter estimation error $\tilde{\vartheta}_I \triangleq \hat{\vartheta}_I^* - \hat{\vartheta}_I$ and the function $\Delta \hat{g}_I \triangleq \hat{g}_I(C_I x_I, z_I, u_I, \hat{\vartheta}_I) - \hat{g}_I(y_I, v_I, u_I, \hat{\vartheta}_I)$, we can write $\Delta g_I = H_I \tilde{\vartheta}_I + \nu_I + \Delta \hat{g}_I$, where $\nu_I \triangleq g_I(C_I x_I, z_I, u_I) - \hat{g}_I(C_I x_I, z_I, u_I, \hat{\vartheta}_I^*)$ is the *Minimum Functional Approximation Error*, with

$$\hat{\vartheta}_I^* \triangleq \arg \min_{\hat{\vartheta}_I \in \Theta_I} \sup_{x_I, z_I, u_I} \|g_I(C_I x_I, z_I, u_I) - \hat{g}_I(C_I x_I, z_I, u_I, \hat{\vartheta}_I)\|$$

is the optimal weight vector. The extended upper bound $\bar{\Delta} g_E$ simply collects the upper bounds of the N subsystems. The threshold in Eq. (13) guarantees that no false-positive alarms will be issued until T_0 because of the uncertainties. In this way, however, the effects of the faults may be “hidden” by the uncertainties in the system dynamics due to the conservativeness of the threshold. This is formalized in the following section in which a distributed detectability sufficient condition will be devised, characterizing the faults that can be detected by the proposed FDI scheme.

3.1 Fault Detectability Analysis

Let us assume that at time $t = T_0$ a fault ϕ occurs in the monolithic system. ϕ_E denotes the extended fault function vector collecting the N subsystems fault functions. After the occurrence of the fault, for $t > T_0$, the state estimation error dynamics becomes

$$\dot{\epsilon}_{x,E} = W [A_{0,E}\epsilon_{x,E} + \Delta f_E + \Delta g_E - L_E\eta_{y,E}(t)] + (1 - e^{-\alpha(t-T_0)})\phi_E \quad (14)$$

and the output estimation error equation for the k -th component is:

$$\begin{aligned} \epsilon_{y,E}^{(k)}(t) &= C_E^{(k)} \epsilon_{x,E}(t) + \eta_{y,E}^{(k)}(t) = \\ C_E^{(k)} \left\{ \int_0^t e^{WA_{0,E}(t-\tau)} [W(\Delta f_E(\tau) + \Delta g_E(\tau) - L_E\eta_{y,E}(\tau)) \right. \\ &\left. + (1 - e^{-\alpha(\tau-T_0)})\phi_E(\tau)d\tau] + e^{WA_{0,E}t}\epsilon_{x,E}(0) \right\} + \eta_{y,E}^{(k)}(t) \quad (15) \end{aligned}$$

Now, we are able to state and prove a sufficient condition for the characterization, in a non-closed form, of a class of faults that can be detected by the proposed FDI methodology.

Theorem 1 (Fault Detectability) *If there exists a time instant $t_1 > T_0$ such that the fault ϕ_E satisfies the inequality*

$$\left| \int_{T_0}^{t_1} C_E^{(k)} e^{WA_{0,E}(t_1-\tau)} (1 - e^{-\alpha(\tau-T_0)}) \phi_E(\tau) d\tau \right| > 2\bar{\epsilon}_{y,E}^{(k)}(t_1) \quad (16)$$

for at least one component $k \in \{1, \dots, p_E\}$, then the fault will be detected at time t_1 , that is $\left| \epsilon_{y,E}^{(k)}(t_1) \right| > \bar{\epsilon}_{y,E}^{(k)}(t_1)$.

Proof 3.2 *At time instant $t_1 > T_0$, the output estimation error can be written as:*

$$\begin{aligned} \epsilon_{y,E}^{(k)}(t_1) &= \int_0^{t_1} C_E^{(k)} e^{WA_{0,E}(t_1-\tau)} [W\Delta f_E(\tau) \\ &+ W\Delta g_E(\tau) - WL_E\eta_{y,E}(\tau)] d\tau + C_E^{(k)} e^{WA_{0,E}t_1} \epsilon_{x,E}(0) \\ &+ \eta_{y,E}^{(k)}(t_1) + \int_{T_0}^{t_1} C_E^{(k)} e^{WA_{0,E}(t_1-\tau)} (1 - e^{-\alpha(\tau-T_0)}) \phi_E(\tau) d\tau \end{aligned}$$

Using the triangle inequality we obtain:

$$\begin{aligned} \left| \epsilon_{y,E}^{(k)}(t_1) \right| &\geq - \left| \int_0^{t_1} C_E^{(k)} e^{WA_{0,E}(t_1-\tau)} [W\Delta f_E(\tau) + \right. \\ &\quad \left. W\Delta g_E(\tau) - WL_E\eta_{y,E}(\tau)] d\tau \right| \\ &\quad + \left| C_E^{(k)} e^{WA_{0,E}t_1} \epsilon_{x,E}(0) + \eta_{y,E}^{(k)}(t_1) \right| \\ &+ \left| \int_{T_0}^{t_1} C_E^{(k)} e^{WA_{0,E}(t_1-\tau)} (1 - e^{-\alpha(\tau-T_0)}) \phi_E(\tau) d\tau \right| \end{aligned}$$

By recalling how the threshold was defined (Eq. 13), it is easy to see that the following inequality holds:

$$\begin{aligned} \left| \epsilon_{y,E}^{(k)}(t_1) \right| &\geq -\bar{\epsilon}_{y,E}^{(k)}(t_1) \\ &+ \left| \int_{T_0}^{t_1} C_E^{(k)} e^{WA_{0,E}(t_1-\tau)} (1 - e^{-\alpha(\tau-T_0)}) \phi_E(\tau) d\tau \right|. \end{aligned}$$

In this way the fault detection condition $\left| \epsilon_{y,E}^{(k)}(t_1) \right| > \bar{\epsilon}_{y,E}^{(k)}(t_1)$ is implied by the theorem hypothesis.

4 Distributed Isolation Architecture

After a fault is detected by any of the N LFDs, the *Global Fault Diagnoser* (GFD) receives the corresponding local fault decision and switches each LFD from fault detection to fault isolation operating mode, stopping the learning of the parameter ϑ_I . Consistently with [34] and

[22], for isolation purposes we assume that the fault function ϕ may belong to a known global fault set \mathcal{F} or be unknown:

$$\mathcal{F} \triangleq \{\phi_1(\mathbf{C}\mathbf{x}, \mathbf{u}), \dots, \phi_{N_{\mathcal{F}}}(\mathbf{C}\mathbf{x}, \mathbf{u})\}.$$

Definition 4.1 *The fault influence set \mathcal{U}_l for the l -th fault function ϕ_l , with $l = 1, \dots, N_{\mathcal{F}}$, is the index set*

$$\mathcal{U}_l \triangleq \{I : \exists t, \exists \mathbf{s}, \mathbf{s} \in \mathcal{I}_I, \phi_l^{(\mathbf{s})}(\mathbf{C}\mathbf{x}, \mathbf{u}) \neq 0\}. \quad (17)$$

In fact, it is possible that not all the subsystems are affected by a given fault function ϕ_l , but only those contained in the corresponding *fault influence set* \mathcal{U}_l . As a consequence, a *local fault set* \mathcal{F}_I can be defined for each subsystem \mathcal{S}_I , collecting the local fault functions $\phi_{I,l}$ such that $I \in \mathcal{U}_l$:

$$\mathcal{F}_I \triangleq \{\phi_{I,1}(C_I x_I, z_I, u_I), \dots, \phi_{I,N_{\mathcal{F}_I}}(C_I x_I, z_I, u_I)\}.$$

It is worth noting that, following Assumption (5), the local fault functions depend only on the local variables x_I, z_I and u_I . It is possible to distinguish two categories of faults, depending upon their topology: *local faults*, whose influence set includes only one subsystem, and *distributed faults*, whose influence set is composed by more than one element.

Besides the FDAE, in the isolation mode each LFD uses other $N_{\mathcal{F}_I}$ estimators called *Fault Isolation Estimators* (FIE), one for each fault in the local fault set \mathcal{F}_I , in order to locally isolate the fault that is acting on the subsystem I . In this way, it is not necessary that the I -th LFD knows the global fault influence set: it is only able to isolate the *local* part of a fault that influences the subsystem \mathcal{S}_I . For each LFD \mathcal{L}_I , with $I = 1, \dots, N$, the generic l -th FIE, with $l \in \{1, \dots, N_{\mathcal{F}_I}\}$, monitors the corresponding fault function $\phi_{I,l}$, belonging to the local fault set \mathcal{F}_I . We assume that each fault function in \mathcal{F}_I can be expressed as:

$$\begin{aligned} \phi_{I,l}(C_I x_I, z_I, u_I) &= [(\vartheta_{I,l,1})^\top H_{I,l,1}(C_I x_I, z_I, u_I), \\ &\quad \dots, (\vartheta_{I,l,n_I})^\top H_{I,l,n_I}(C_I x_I, z_I, u_I)]^\top, \quad (18) \end{aligned}$$

where $H_{I,l,k} : \mathbb{R}^{p_I} \times \mathbb{R}^{q_I} \times \mathbb{R}^{m_I} \mapsto \mathbb{R}^{q_{I,l,k}}$, with $k \in \{1, \dots, n_I\}$, $l \in \{1, \dots, N_{\mathcal{F}_I}\}$, are the known functions describing the functional structure of the fault and $\vartheta_{I,l,k} \in \Theta_{I,l,k} \subset \mathbb{R}^{q_{I,l,k}}$ are the unknown parameter vectors providing its ‘‘magnitude’’, where the parameter domains $\Theta_{I,l,k}$ are assumed to be origin-centered hyper-spheres with radius $M_{\Theta_{I,l,k}}$, without much loss of generality.

After $t = T_d$, the generic l -th FIE estimator is activated and monitors its subsystem, computing a local state estimate $\hat{x}_{I,l}$ and a local output estimate $\hat{y}_{I,l}$. The difference between the estimate $\hat{y}_{I,l}$ and the measurements y_I

is the estimation error $\epsilon_{y,I,l} \triangleq y_I - \hat{y}_{I,l}$, used as a residual and compared, component by component, to an appropriate isolation threshold $\bar{\epsilon}_{y,I,l} \in \mathbb{R}_+^{p_I}$. The condition

$$|\epsilon_{y,I,l}^{(k)}(t)| \leq \bar{\epsilon}_{y,I,l}^{(k)}(t) \quad \forall k = 1, \dots, p_I \quad (19)$$

is associated to the l -th fault hypothesis

$\mathcal{H}_{I,l}$: “The subsystem \mathcal{S}_I is affected by the l -th fault” .
(20)

Should the condition be violated at some time instant t , then the hypothesis $\mathcal{H}_{I,l}$ will be falsified excluding the fault $\phi_{I,l}$ and the *local fault isolation signature* $\mathcal{S}_{I,l}$ will become non-empty:

Definition 4.2 The l -th local isolation signature shown by the subsystem \mathcal{S}_I , $I \in \{1, \dots, N\}$, $l \in \{1, \dots, N_{\mathcal{F}_I}\}$ at time $t > 0$ is the index set

$$\mathcal{S}_{I,l}(t) \triangleq \{k \in \{1, \dots, p_I\} : \exists t_1, t \geq t_1 > 0 \text{ such that } |\epsilon_{y,I,l}^{(k)}(t)| > \bar{\epsilon}_{y,I,l}^{(k)}(t_1)\}. \quad (21)$$

Definition 4.3 The l -th fault exclusion time $T_{e,I,l}$ is defined as $T_{e,I,l} \triangleq \min\{t : \mathcal{S}_{I,l}(t) \neq \emptyset\}$.

Definition 4.4 A fault $\phi_{I,\rho} \in \mathcal{F}_I$ is locally isolated at time t iff $\forall l, l \in \{1, \dots, N_{\mathcal{F}_I}\} \setminus \rho$, $\mathcal{S}_{I,l}(t) \neq \emptyset$ and $\mathcal{S}_{I,\rho}(t) = \emptyset$. Furthermore $T_{locisol,I,\rho} \triangleq \max\{T_{e,I,l}, l \in \{1, \dots, N_{\mathcal{F}_I}\} \setminus \rho\}$ is the local fault isolation time.

The aim of the isolation task is to exclude every but one fault: we can conclude that a specific fault actually occurred if we assume that only faults belonging to the set \mathcal{F}_I may occur.

Now, a detailed description of the FIEs is given. After the fault $\phi(t)$ has occurred, the s_I -th component of the I -th local state equation becomes

$$\begin{aligned} \dot{x}_I^{(s_I)} &= A_I^{(s_I)} x_I + f_I^{(s_I)}(x_I, u_I) \\ &+ g_I^{(s_I)}(C_I x_I, z_I, u_I) + \beta(t - T_0) \phi^{(s)}(C x, u). \end{aligned}$$

The l -th FIE computes a local estimate, that, in the case of non-shared state variables, can be defined as:

$$\begin{cases} \dot{\hat{x}}_{I,l} = A_I \hat{x}_{I,l} + f_I(\hat{x}_{I,l}, u_I) + \hat{g}_I(y_I, u_I, v_I, \hat{\vartheta}_{I,0}) \\ \quad + L_I(y_I - \hat{y}_{I,l}) + \hat{\phi}_{I,l}(y_I, v_I, u_I, \hat{\vartheta}_{I,l}) \\ \hat{y}_{I,l} = C_I \hat{x}_{I,l}, \end{cases} \quad (22)$$

where $L_I \in \mathbb{R}^{n_I \times p_I}$ is the local output error gain, $\hat{\phi}_{I,l}^{(s_I)}(y_I, v_I, u_I, \hat{\vartheta}_{I,l}) \triangleq (\hat{\vartheta}_{I,l, s_I})^\top H_{I,l, s_I}(y_I, v_I, u_I)$ is the s_I -th component of a linearly-parameterized function that learns the structure of the l -th fault function $\phi_{I,l}$, where the vector $\hat{\vartheta}_{I,l} \triangleq \text{col}(\hat{\vartheta}_{I,l,k}, k \in \{1, \dots, n_I\})$

contains its adjustable parameters. We developed the following learning law, for every $I = 1, \dots, N$, $l = 1, \dots, N_{\mathcal{F}_I}$:

$$\dot{\hat{\vartheta}}_{I,l} = P_{\hat{\vartheta}_{I,l}} [\Gamma_{I,l} H_{I,l}^\top W_{I,l}^\top C_I^\top \epsilon_{y,I,l}] \quad (23)$$

$$H_{I,l}^\top = \partial \hat{\phi}_{I,l} / \partial \hat{\vartheta}_{I,l}.$$

The dynamics of the l -th FIE estimator for the most general case of a distributed fault can be defined as

$$\begin{aligned} \dot{\hat{x}}_{I,l}^{(s_I)} &= \sum_{J \in \mathcal{O}_s} W_s^{(I,J)} \left[A_J^{(s_J)} \hat{x}_{J,l} + f_J^{(s_J)}(\hat{x}_{J,l}(t), u_J) \right. \\ &+ \hat{g}_J^{(s_J)}(y_J, u_J, v_J, \hat{\vartheta}_{J,0}) + L_J^{(s_J)}(y_{J,l} - \hat{y}_{J,l}) \\ &\left. + \hat{\phi}_{J,l}^{(s_J)}(y_J, v_J, u_J, \hat{\vartheta}_{J,l}) \right] \quad (24) \end{aligned}$$

The i -th component of the estimation error, in the non-shared case, is

$$\begin{aligned} \dot{\epsilon}_{x,I,l}^{(i)} &= A_{0,I}^{(i)} \epsilon_{x,I,l} + \Delta f_I^{(i)} + \Delta g_I^{(i)} - L_I^{(i)} \eta_{y,I} \\ &+ (1 - e^{-\alpha(t-T_0)}) \phi^{(i)} - \hat{\phi}_{J,l}^{(i)}(y_J, v_J, u_J, \hat{\vartheta}_{J,l}), \quad (25) \end{aligned}$$

On the other hand, the dynamics of the state estimation error component for shared variables can be described as:

$$\begin{aligned} \dot{\epsilon}_{x,I,l}^{(s_I)} &= A_I^{(s_I)} x_I + f_I^{(s_I)}(x_I, u_I) + g_I^{(s_I)}(C_I x_I, u_I, z_I) + \\ &(1 - e^{-\alpha(t-T_0)}) \phi^{(s)} - \sum_{J \in \mathcal{O}_s} W_s^{(I,J)} \left[A_J^{(s_J)} \hat{x}_{J,l} \right. \\ &+ f_J^{(s_J)}(\hat{x}_{J,l}, u_J) + \hat{g}_J^{(s_J)}(y_J, u_J, v_J, \hat{\vartheta}_{J,0}) \\ &\left. + L_J^{(s_J)}(y_J - \hat{y}_{J,l}) + \hat{\phi}_{J,l}^{(s_J)} \right]. \end{aligned}$$

When we consider a matched fault

$$\phi^{(s)} = \phi_{J,l}^{(s_J)}(x_J, z_J, u_J, \vartheta_{J,l}), \quad \forall J \in \mathcal{O}_s,$$

the error dynamics can then be rewritten as:

$$\begin{aligned} \dot{\epsilon}_{x,I,l}^{(s_I)} &= \sum_{J \in \mathcal{O}_s} W_s^{(I,J)} \left[A_{0,J}^{(s_J)} \epsilon_{x,J,l} + \Delta f_J^{(s_J)} \right. \\ &\left. + \Delta g_J^{(s_J)} - L_J^{(s_J)} \eta_{y,J} + \Delta \phi_{J,l}^{(s_J)} \right] \quad (26) \end{aligned}$$

where

$$\begin{aligned} \Delta \phi_{J,l}^{(s_J)} &\triangleq (1 - e^{-\alpha(t-T_0)}) \phi^{(s)} - \hat{\phi}_{J,l}^{(s_J)} = (1 - e^{-\alpha(t-T_0)}) \\ &(H_{J,l, s_J}^\top(t)^\top \vartheta_{J,l, s_J} + \Delta H_{J,l, s_J}^\top \vartheta_{J,l, s_J}) - H_{J,l, s_J}(t)^\top \hat{\vartheta}_{J,l, s_J} \end{aligned}$$

with $\Delta H_{J,l,s_J}^\top \triangleq H_{J,l,s_J}(x_J, z_J, u_J) - H_{J,l,s_J}(y_J, v_J, u_J)$. It can be rewritten as

$$\begin{aligned} \Delta \phi_{J,l}^{(s_J)} &= (1 - e^{-\alpha(t-T_0)}) H_{J,l,s_J}(t)^\top \tilde{\vartheta}_{J,l,s_J} \\ &\quad + (1 - e^{-\alpha(t-T_0)}) \Delta H_{J,l,s_J}(t)^\top \vartheta_{J,l,s_J} \\ &\quad - e^{-\alpha(t-T_0)} H_{J,l,s_J}(t)^\top \hat{\vartheta}_{J,l,s_J} \end{aligned}$$

if we introduce the parameter estimation errors $\tilde{\vartheta}_{J,l,s_J} \triangleq \vartheta_{J,l,s_J} - \hat{\vartheta}_{J,l,s_J}$. Using the general formulation, we can express the dynamics of the estimation error in the case of a matched fault, as:

$$\begin{aligned} \dot{\epsilon}_{x,E,l} &= W [A_{0,E} \epsilon_{x,E,l} + \Delta f_E + \Delta g_E - L_E \eta_{y,E} \\ &\quad + \Delta \phi_{E,l}]. \end{aligned}$$

We can now compute the state estimation error solution:

$$\begin{aligned} \epsilon_{x,E,l}(t) &= \int_{T_d}^t e^{W A_{0,E}(t-\tau)} [W \Delta f_E(\tau) + W \Delta g_E(\tau) \\ &\quad - W L_E \eta_{y,E}(\tau) + W \Delta \phi_{E,l}(\tau) d\tau] \\ &\quad + e^{W A_{0,E}(t-T_d)} \epsilon_{x,E,l}(T_d) \quad (27) \end{aligned}$$

The output estimation error in the case of a matched fault can be written componentwise as: $\epsilon_{y,E,l}^{(k)} \triangleq C_E^{(k)} \epsilon_{x,E,l} + \eta_{y,E}^{(k)}$, for all $k = 1, \dots, p_E$. It can be bounded by:

$$\begin{aligned} \left| \epsilon_{y,E,l}^{(k)}(t) \right| &\leq \left| C_E^{(k)} \left\{ \int_{T_d}^t e^{W A_{0,E}(t-\tau)} [W (\Delta f_E(\tau) \right. \right. \\ &\quad \left. \left. + \Delta g_E(\tau) - L_E \eta_{y,E}(\tau) + \Delta \phi_{E,l}(\tau)) d\tau \right. \right. \\ &\quad \left. \left. + e^{W A_{0,E}(t-T_d)} \epsilon_{x,E,l}(T_d) \right\} \right| + \left| \eta_{y,E}^{(k)}(t) \right| \\ &\leq \left| C_E^{(k)} \right| \left\{ \int_{T_d}^t \left\| e^{W A_{0,E}(t-\tau)} \right\| \left\| W [\bar{\Delta} f_E(\tau) \right. \right. \\ &\quad \left. \left. + \bar{\Delta} g_E(\tau) + |L_E| \bar{\eta}_{y,E}(\tau) + \bar{\Delta} \phi_{E,l}(\tau) d\tau \right. \right. \\ &\quad \left. \left. + \left\| e^{W A_{0,E}(t-T_d)} \right\| \bar{\epsilon}_{x,E,l}(T_d) \right\} + \bar{\eta}_{y,E}^{(k)}(t) \quad (28) \end{aligned}$$

where

$$\begin{aligned} \bar{\Delta} \phi_{E,l} &= \text{col}(\|H_{I,l,s_I}(t)\| \kappa_{I,l,s_I}(\hat{\vartheta}_{I,l,s_I}) \\ &\quad + \bar{\Delta} H_{I,l,s_I}(t) \bar{\vartheta}_{I,l,s_I} - e^{-\bar{\alpha}(t-T_d)} \|H_{I,l,s_I}(t)\| \|\hat{\vartheta}_{I,l,s_I}\|, \\ &\quad s_I = 1, \dots, n_I, I = 1, \dots, N). \end{aligned}$$

where $\kappa_{I,l}(\hat{\vartheta}_{I,l}) \geq \|\hat{\vartheta}_{I,l}\|$. The threshold (28) can be computed in a distributed way and, because of the way it is defined, guarantees that no matched fault will be excluded due to the presence of uncertainties or to the effect of the parameter estimation error $\hat{\vartheta}_{I,l}$.

4.1 Fault isolability analysis

We now consider the case of a non-matched fault $\phi_I^{(s_I)}(x_I, z_I, u_I) = \phi_{I,\rho}^{(s_I)}(x_I, z_I, u_I, \vartheta_{I,\rho})$, with $\rho \neq l$. In this case, the dynamics of the shared s_I -component of the estimation error for the l -th FIE of the I -th LFD can be written as

$$\begin{aligned} \dot{\epsilon}_{x,I,l} &= \sum_{J \in \mathcal{O}_s} W_s^{(I,J)} [A_{0,J}^{(s_J)} \epsilon_{x,J,l} + \Delta f_J^{(s_J)} + \Delta g_J^{(s_J)} \\ &\quad - L_J^{(s_J)} \eta_{y,J} + (1 - e^{-\alpha(t-T_0)}) \phi_{I,\rho}^{(s_I)}(x_I, z_I, u_I, \vartheta_{I,\rho}) \\ &\quad - \hat{\phi}_{J,l}^{(s_J)}(y_J, v_J, u_J, \hat{\vartheta}_{J,l})]. \end{aligned}$$

and considering the vector $\epsilon_{x,E,l}$:

$$\begin{aligned} \dot{\epsilon}_{x,E,l} &= W [A_{0,E} \epsilon_{x,E,l} + \Delta f_E + \Delta g_E - L_E \eta_{y,E} \\ &\quad + \Delta_l \phi_{E,\rho}], \end{aligned}$$

where the *mismatch vector* is introduced

$$\begin{aligned} \Delta_l \phi_{E,\rho} &\triangleq \text{col}((1 - e^{-\alpha(t-T_0)}) \phi_{I,\rho}^{(s_I)} - \hat{\phi}_{I,l}^{(s_I)}), \\ &\quad s_I = 1, \dots, n_I, I = 1, \dots, N). \end{aligned}$$

The solution can then be written as

$$\begin{aligned} \epsilon_{x,E,l}(t) &= \int_{T_d}^t e^{W A_{0,E}(t-\tau)} W [\Delta f_E(\tau) + \Delta g_E(\tau) \\ &\quad - L_E \eta_{y,E}(\tau) + \Delta_l \phi_{E,\rho}(\tau) d\tau] + e^{W A_{0,E}(t-T_d)} \epsilon_{x,E,l}(T_d) \end{aligned}$$

and so, the output residual can be expressed componentwise by the following equation:

$$\begin{aligned} \epsilon_{y,E,l}^{(k)}(t) &= \eta_{y,E}^{(k)}(t) \\ &\quad + C_E^{(k)} \left\{ \int_{T_d}^t e^{W A_{0,E}(t-\tau)} W [\Delta f_E(\tau) + \Delta g_E(\tau) \right. \\ &\quad \left. - L_E \eta_{y,E}(\tau) + \Delta_l \phi_{E,\rho}(\tau) d\tau] + e^{W A_{0,E}(t-T_d)} \epsilon_{x,E,l}(T_d) \right\} \end{aligned}$$

At this point, a sufficient condition for fault isolability can be proved.

Theorem 2 (Fault Isolability) *Given a fault $\phi_{I,\rho} \in \mathcal{F}_I$, if for each $l \in \{1, \dots, N_{\mathcal{F}_I}\} \setminus \rho$, the following inequality holds*

$$\begin{aligned} \left| \int_{T_d}^{T_l} C_E^{(k)} e^{W A_{0,E}(T_l-\tau)} \Delta_l \phi_{E,\rho}(\tau) d\tau \right| &> \bar{\epsilon}_{y,E,l}^{(k)}(T_l) \\ &+ \left| C_E^{(k)} \right| \left\{ \int_{T_d}^{T_l} \left\| e^{W A_{0,E}(T_l-\tau)} \right\| \left\| W [\bar{\Delta} f_E(\tau) + \bar{\Delta} g_E(\tau) \right. \right. \\ &\quad \left. \left. + |L_E| \bar{\eta}_{y,E}(\tau) d\tau] + \left\| e^{W A_{0,E}(T_l-T_d)} \right\| \bar{\epsilon}_{x,E,l}(T_d) \right\} + \bar{\eta}_{y,E}^{(k)}(T_l) \end{aligned}$$

at some time instant $T_l > T_d$, for some $k \in \{1, \dots, p_I\}$, then the ρ -th fault will be isolated. The local isolation time is upper-bounded by $\max_{l \in \{1, \dots, N_{\mathcal{F}_I}\} \setminus \rho} (T_l)$.

Proof 4.1 After fault detection, for $T_l > T_d$, we can bound the absolute value of the k -th component of the l -th FIE estimation error using the triangle inequality:

$$\begin{aligned} |\epsilon_{y,E,l}^{(k)}(T_l)| &\geq \left| \int_{T_d}^{T_l} C_E^{(k)} e^{WA_{0,E}(T_l-\tau)} \Delta_l \phi_{E,\rho}(\tau) d\tau \right| \\ &- \left| C_E^{(k)} \int_{T_d}^{T_l} e^{WA_{0,E}(T_l-\tau)} [W \Delta f_E(\tau) + W \Delta g_E(\tau) \right. \\ &\left. - W L_E \eta_{y,E}(\tau) d\tau \right] \right| - \left| C_E^{(k)} e^{WA_{0,E}(T_l-T_d)} \epsilon_{x,E,l}(T_d) \right| \\ &- \left| \eta_{y,E}^{(k)}(T_l) \right| \end{aligned}$$

We want the inequality $|\epsilon_{y,E,l}^{(k)}(T_l)| > \bar{\epsilon}_{y,E,l}^{(k)}(T_l)$ to be satisfied in order to exclude the l -th fault. This results in the following inequality

$$\begin{aligned} \left| \int_{T_d}^{T_l} C_E^{(k)} e^{WA_{0,E}(T_l-\tau)} \Delta_l \phi_{E,\rho}(\tau) d\tau \right| &> \bar{\epsilon}_{y,E,l}^{(k)}(T_l) \\ + \left| C_E^{(k)} \int_{T_d}^{T_l} e^{WA_{0,E}(T_l-\tau)} [W \Delta f_E(\tau) + W \Delta g_E(\tau) \right. \\ &\left. - W L_E \eta_{y,E}(\tau) d\tau \right] \right| + \left| C_E^{(k)} e^{WA_{0,E}(T_l-T_d)} \epsilon_{x,E,l}(T_d) \right| \\ &+ \left| \eta_{y,E}^{(k)}(T_l) \right| \end{aligned}$$

which is implied by the inequality in the hypothesis of the theorem. This fault is isolated in the sense of Definition 4.4 if the inequality holds for every fault function of \mathcal{F}_I but the ρ -th.

4.2 Global fault isolation logic

In this subsection, we analyze the global fault isolation logic. The GFD, which is assumed to know both the global fault set \mathcal{F} and the fault influence sets of all the global fault functions, receives the local fault decisions d_I^{FD} from each LFD and determines which one of the faults, if any, in the global set \mathcal{F} affects the system \mathcal{S} . As proposed in [22], it is important to make a distinction between local and distributed faults. For a local fault, it is sufficient that the corresponding LFD excludes every but that fault for concluding that it is isolated. Instead, in the case of distributed faults, the isolation requires that all the LFDs in the influence set of that fault, exclude all other faults. In both cases, it holds:

Definition 4.5 A fault $\phi_l \in \mathcal{F}$ is globally isolated if for each J -th LFD in the fault influence set U_l , the corresponding local functions $\phi_{J,l,J}$ have been isolated, with $J \in \mathcal{U}_l$. Furthermore $T_{isol,l} \triangleq \max\{T_{locisol,J,l,J}, J \in \mathcal{U}_l\}$ is the global fault isolation time.

5 Concluding Remarks

In this paper, the problem of distributed fault detection and isolation with partial state measurement was addressed in the case of continuous-time systems, by extending a diagnosis architecture developed by the same authors in [28] for discrete-time systems. Future research efforts will be devoted to investigate the conservativeness of the detectability and isolability conditions and to show the effectiveness of the proposed technique by validation on practically-relevant distributed use-cases, both in simulation and in actual experiments.

References

- [1] M. E. J. Newman, A. L. Barabasi, D. J. Watts, The structure and dynamics of networks, Princeton University Press, 2006.
- [2] J. Baillieul, P. J. Antsaklis, Control and communication challenges in networked real-time systems, Proc. IEEE 95 (1) (2007) 9–28.
- [3] C. T. Abdallah, H. G. Tanner, Complex networked control systems: introduction to the special section, Control Systems Magazine, IEEE 27 (4) (2007) 30–32.
- [4] R. D’Andrea, G. E. Dullerud, Distributed control design for spatially interconnected systems, IEEE Trans. Autom. Control 48 (9) (2003) 1478–1495.
- [5] J. Cortes, S. Martinez, T. Karatas, F. Bullo, Coverage control for mobile sensing networks, Robotics and Automation 20 (2) (2004) 243–255.
- [6] R. Olfati-Saber, Flocking for multi-agent dynamic systems: algorithms and theory, IEEE Trans. Autom. Control 51 (3) (2006) 401–420.
- [7] N. Sandell, P. Varaiya, M. Athans, M. Safonov, Survey of decentralized control methods for large scale systems, IEEE Trans. Autom. Control 23 (2) (1978) 108–128.
- [8] P. Baroni, G. Lamperti, P. Pogliano, M. Zanella, Diagnosis of large active systems, Artif. Intell. 110 (1) (1999) 135–189.
- [9] F. P. Preparata, G. Metzger, R. T. Chien, On the connection assignment problem of diagnosable systems, IEEE Trans. Electron. Comput. EC-16 (6) (1967) 848–854.
- [10] A. Fagiolini, G. Valenti, L. Pallottino, G. Dini, Decentralized intrusion detection for secure cooperative multi-agent systems, in: Proc. 46th IEEE Conf. on Decision and Control, 2007, pp. 1553–1558.
- [11] C. Edwards, L. M. Fridman, M. W. Thein, Fault reconstruction in a leader/follower spacecraft system using higher order sliding mode observers, in: Proc. American Control Conference, New York, USA, 2007, pp. 408–413.
- [12] A. Teixeira, S. Amin, H. Sandberg, K. H. Johansson, S. S. Sastry, Cyber security analysis of state estimators in electric power systems, in: 49th IEEE Conf. on Decision and Control, 2010, pp. 5991–5998.

- [13] S. Jafari, A. Ajorlou, A. G. Aghdam, S. Tafazoli, On the structural controllability of multi-agent systems subject to failure: A graph-theoretic approach, *Proc. 49th IEEE Conf. on Decision and Control* (2010) 4565–4570.
- [14] M. Daigle, X. Koutsoukos, G. Biswas, Distributed diagnosis in formations of mobile robots, *IEEE Trans. Robot.* 23 (2007) 353–369.
- [15] N. Lechevin, C. A. Rabbath, Decentralized detection of a class of non-abrupt faults with application to formations of unmanned airships, *Control Systems Technology, IEEE Transactions on* 17 (2) (2009) 484–493.
- [16] W. Li, W. H. Gui, Y. F. Xie, S. X. Ding, Decentralized fault detection system design for large-scale interconnected systems, in: *Proc. of 7th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes (SafeProcess 2009)*, Barcelona, Spain, 2009, pp. 816–821.
- [17] S. Stankovič, N. Ilić, Z. Djurovič, M. Stankovič, K. H. Johansson, Consensus based overlapping decentralized fault detection and isolation, *Conf. on Control and Fault-Tolerant Systems*, (2010) 570–575.
- [18] X. Zhang, Decentralized fault detection for a class of large-scale nonlinear uncertain systems, in: *Proc. 2010 American Control Conference*, 2010, pp. 5650–5655.
- [19] V. Reppa, M. Polycarpou, C. Panayiotou, Distributed sensor fault detection and isolation for nonlinear uncertain systems, in: *Proc. of Safeprocess Conference*, 2012, pp. 1077–1082.
- [20] F. Boem, R. M. G. Ferrari, T. Parisini, M. M. Polycarpou, Distributed fault diagnosis for input-output continuous-time nonlinear systems, in: *Proc. of Safeprocess Conference*, 2012, pp. 1089–1094.
- [21] F. Boem, R. M. G. Ferrari, T. Parisini, Distributed fault detection and isolation of continuous-time nonlinear systems, *Europ. J. of Control* (5-6) (2011) 603–620.
- [22] R. M. G. Ferrari, T. Parisini, M. M. Polycarpou, Distributed fault detection and isolation of large-scale discrete-time nonlinear systems: an adaptive approximation approach, *IEEE Trans. Autom. Control* 57 (2) (2012) 275–290.
- [23] A. T. Vemuri, M. M. Polycarpou, Robust nonlinear fault diagnosis in input-output systems, *International Journal of Control* 68 (2) (1997) 343–360.
- [24] Q. Zhang, M. Basseville, A. Benveniste, Fault detection and isolation in nonlinear dynamic systems: A combined input–output and local approach, *Automatica* 34 (11) (1998) 1359–1373.
- [25] X. Zhang, M. M. Polycarpou, T. Parisini, Robust fault isolation for a class of non-linear input–output systems, *Internat. J. Control* 74 (13) (2001) 1295–1310.
- [26] K. Subbarao, A. Vemuri, Fault isolation using extrinsic curvature for multi-input-multi-output systems with nonlinear fault models, in: *Proc. American Control Conference*, 2007, pp. 3246–3251.
- [27] Z. Zhang, I. M. Jaimoukha, Fault detection and isolation for linear discrete-time systems using input/output measurement analysis, in: *Proc. 48th IEEE Conf. on Decision and Control and 28th Chinese Control Conf.*, 2009, pp. 4908–4913.
- [28] F. Boem, R. M. G. Ferrari, T. Parisini, M. M. Polycarpou, A distributed fault detection methodology for a class of large-scale uncertain input-output discrete-time nonlinear systems, in: *Proc. 50th IEEE Conf. on Decision and Control and European Control Conf.*, 2011.
- [29] X. Zhang, Q. Zhang, Distributed fault diagnosis in a class of interconnected nonlinear uncertain systems, *International Journal of Control* 85 (11) (2012) 1644–1662.
- [30] D. D. Šiljak, *Large-Scale Dynamic Systems: Stability and Structure*, North Holland, New York, 1978.
- [31] L. Xiao, S. Boyd, S. Kim, Distributed average consensus with least-mean-square deviation, *J. Par. Dist. Comput.* 67 (1) (2007) 33–46.
- [32] C. D. Meyer (Ed.), *Matrix analysis and applied linear algebra*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2000.
- [33] M. M. Polycarpou, On-line approximators for nonlinear system identification: a unified approach, in: X. Lendes (Ed.), *Control and Dynamic Systems: Neural Network Systems Techniques and Applications*, Vol. 7, Academic, New York, 1998, pp. 191–230.
- [34] X. Zhang, M. M. Polycarpou, T. Parisini, A robust detection and isolation scheme for abrupt and incipient faults in nonlinear systems, *IEEE Trans. Autom. Control* 47 (4) (2002) 576–593.