

Distributed Fault Diagnosis for Process and Sensor Faults in a Class of Interconnected Input-Output Nonlinear Discrete-Time Systems

Christodoulos Keliris^a, Marios M. Polycarpou^a, Thomas Parisini^{b*}

^a*KIOS Research Center for Intelligent Systems and Networks, Dept. of Electrical and Computer Engineering, University of Cyprus, Nicosia 1678, Cyprus*

^b*Dept. of Electrical and Electronic Engineering, Imperial College London, UK, and Dept. of Engineering and Architecture, University of Trieste, Italy*

(Received 00 Month 20XX; accepted 00 Month 20XX)

This paper presents a distributed fault diagnosis scheme able to deal with process and sensor faults in an integrated way for a class of interconnected input-output nonlinear uncertain discrete-time systems. A robust distributed fault detection scheme is designed, where each interconnected subsystem is monitored by its respective fault detection agent, and according to the decisions of these agents, further information regarding the type of the fault can be deduced. As it is shown, a process fault occurring in one subsystem can only be detected by its corresponding detection agent whereas a sensor fault in a subsystem can be detected by either its corresponding detection agent or the detection agent of another subsystem that is affected by the subsystem where the sensor fault occurred. This discriminating factor is exploited for the derivation of a high-level isolation scheme. Moreover, process and sensor fault detectability conditions characterizing quantitatively the class of detectable faults are derived. Finally, a simulation example is used to illustrate the effectiveness of the proposed distributed fault detection scheme.

Keywords: nonlinear systems; fault diagnosis; filtering; fault propagation; process and sensor faults

1. Introduction

In many applications involving large-scale systems, collaboration and information exchange among several subsystems is of crucial importance. Examples include power systems, communication networks and water systems. If the problem of real-time monitoring is not properly addressed, the operation of such systems may create life-threatening situations and cause significant economic damage. Therefore, the development of robust fault detection approaches for identifying promptly any abnormal system behavior is a primary task for achieving safe and reliable system operation.

The area of fault diagnosis is at the forefront of the technological evolution for several decades. Many survey papers (Frank, 1990; Venkatasubramanian, Rengaswamy, Yin, & Kavuri, 2003) and books (Blanke, Kinnaert, Lunze, & Staroswiecki, 2010; Chen & Patton, 1999; R. Patton & Frank, 1989) exist on the topic. The problem of fault detection and isolation (FDI) for linear systems is well investigated and the interested reader is directed to the aforementioned books and survey papers. In the case of FDI for nonlinear systems, until recently, centralized fault diagnosis approaches were the main topic of investigation and a variety of methods were employed; i.e. by using nonlinear observer design (Hammouri, Kinnaert, & El Yaagoubi, 1999; Rajamani & Ganguli, 2004), adaptive estimation methods (Reppa, Polycarpou, & Panayiotou, 2014; X. Zhang, Polycarpou, & Parisini,

*Corresponding author. Email: t.parisini@gmail.com

2002), change detection methods (Basseville & Nikiforov, 1993; Q. Zhang, Basseville, & Benveniste, 1998) and differential-geometric approaches (De Persis & Isidori, 2002). In the last years though, due to advances in computing and communications mainly, the focus of the research activities is directed mostly towards the development of hierarchical (Klinkhio & Patton, 2009; R. J. Patton et al., 2007), decentralized (Ferdowsi, Raja, & Jagannathan, 2012; Léchevin & Rabbath, 2009; Stankovic, Ilic, Djurovic, Stankovic, & Johansson, 2010; Wei, Gui, Xie, & Ding, 2009; X. Zhang, Polycarpou, & Parisini, 2009) and distributed (Boem, Ferrari, & Parisini, 2011; Ferrari, Parisini, & Polycarpou, 2012; Keliris, Polycarpou, & Parisini, 2013a; Yan, Tian, & Shi, 2008) schemes for process or sensor faults. In our earlier work Keliris et al. (2013a), a distributed fault detection approach under continuous time and full state measurements was presented for the case of process faults along with a thorough detectability analysis. The approach made use of filtering in order to dampen the effects of measurement noise and aid in the derivation of tight detection thresholds. In Keliris, Polycarpou, and Parisini (2013b), the approach given in Keliris et al. (2013a) was extended by relaxing the assumption of the availability of all the state measurements and, the proposed approach dealt only with process faults under continuous time. In this work, we maintain the use of filtering and we investigate further the properties of the filtering approach under discrete time, by considering the input-output case and, by dealing with process and sensor faults in an integrated way.

In many cases, the architecture of the underlying subsystems that are inherently decentralized or distributed makes the development of a distributed FDI framework a necessity. For instance, many factors contribute to the need for a distributed FDI formulation such as the large scale nature of the system to be monitored, its spatial distribution, the inability to access centrally certain parts of the system. As a result, local diagnosis should be performed, however the increasing complexity of large scale interconnected systems creates additional difficulties to the fault diagnosis problem, especially with issues related to fault propagation where, a fault that occurs in one subsystem appears and affects neighboring subsystems. Therefore, there is the need to understand better the fault propagation issues involved and the behavior of the system and the corresponding FDI scheme. This paper, contributes towards these directions by proposing a distributed fault detection scheme and addressing fault propagation among interconnected subsystems.

In the research literature there is the tendency to deal with the problem of fault diagnosis for process and sensor faults separately, something that poses a significant limitation for real world applications. For example, in the problem of fault diagnosis for process faults the sensors are assumed healthy. However, apart from erroneous detection results, a faulty sensor may also lead to degraded tracking or regulation performance or even endanger the stability of the control system. Acknowledging, that sensors are prone to faults and utilizing sensor validation approaches is crucial to the overall system stability and reliability. Similarly, in the problem of fault diagnosis for sensor faults, typically it is assumed that there are no process faults. Obviously, dealing with the process and sensor fault problem separately incurs the danger of false alarms due to monitoring the specific fault type and ignoring the other, something that results in unnecessary component replacement and increased maintenance costs.

The research conducted on the fault diagnosis problem that deals simultaneously with process and sensor faults is limited (Dunia & Joe Qin, 1998; Salahshoor, Mosallaei, & Bayat, 2008; Q. Zhang & Zhang, 2012; X. Zhang, Polycarpou, & Parisini, 2008). Some earlier results dealing jointly with actuator and sensor faults can also be found in Kinnaert and Peng (1995); Massoumnia and Vander Velder (1988); Viswanadham and Srichander (1987). In Salahshoor et al. (2008) the sensor and process fault detection problem is addressed using multi sensor data fusion techniques based on the adaptive extended Kalman filter algorithm, whereas in Dunia and Joe Qin (1998) a unified framework for dealing with joint diagnosis of process and sensor faults is proposed along with fault identification and reconstruction via principal component analysis. In the context of analytical redundancy methods, X. Zhang et al. (2008) develop a fault isolation approach to determine which process or sensor fault, among two respective fault classes, has occurred. Based on the assumption that only a single fault occurs (either a process or a single sensor), adaptive approximation methods are used in order to build a fault detection estimator and suitable fault isolation estimators that correspond to the

process and sensor faults that are able to determine which fault has occurred. In Talebi, Khorasani, and Tafazoli (2009) a recurrent neural-network based fault detection scheme for nonlinear systems is proposed, which employs two nonlinear-in-parameters neural networks to isolate actuator and sensor faults; the fault determined when the output of one of the neural networks produces a non-zero output indicating the faulty condition. In Thumati and Halligan (2013) a nonlinear observer-based fault diagnostics scheme, dealing with process and sensor faults, for nonlinear systems in discrete time is proposed. The scheme consists of an artificial immune system as an online approximator, which identifies the fault type by monitoring the outputs' magnitude of the two online approximators (state and output) as in Talebi et al. (2009). In Q. Zhang and Zhang (2012) a distributed detection scheme for process and sensor faults for a class of input-output interconnected systems under continuous time is proposed, but the estimator design is conducted under some potentially restrictive conditions and, deals only with the fault detectability issue. This paper contributes to the limited research available, by developing a distributed fault detection scheme for process and sensor faults and investigating the propagation of the fault effects to neighboring subsystems.

The primary objective and main contribution of this paper is the derivation of a distributed fault diagnosis approach dealing with process and sensor faults in an integrated way, utilizing a specifically designed scheme that encompasses important characteristics regarding fault propagation among subsystems. Detectability conditions for process and sensor faults are derived, characterizing quantitatively the class of detectable faults. The scheme is comprised of a set of interacting fault detection agents, in which each subsystem is monitored by its respective detection agent. As shown, a process fault that occurs in a subsystem can only be detected by its respective detection agent, whilst a sensor fault that occurs in a subsystem may also be detected by the detection agents of neighboring subsystems that are affected by the subsystem where the sensor fault occurred. This differentiating element is exploited in order to derive a high-level fault isolation scheme, able to provide information regarding the type and location of the fault that has occurred. Therefore, the proposed distributed fault detection approach encompasses significant benefits for the fault isolation task which can be exploited by a more sophisticated isolation scheme to pinpoint the exact fault that occurred.

The paper is organized as follows: the problem formulation is given in Section 2 and the detailed design of the distributed fault detection scheme is presented in Section 3. In Section 4, the detectability conditions that characterize the class of detectable process and sensor faults are derived. In Section 5, fault propagation issues are investigated, and in Section 6, a high-level fault isolation scheme is proposed. In Section 7, a simulation example demonstrating the effectiveness of the scheme is presented and, finally, in Section 8 some concluding remarks are given.

2. Problem formulation

We consider an interconnected nonlinear dynamic system comprised of N subsystems Σ_I , $I \in \{1, \dots, N\}$. The discrete-time dynamics of each subsystem is described by:

$$\Sigma_I : \begin{cases} x_I(t+1) = A_I x_I(t) + g_I(y_I^0(t), \bar{y}_I^0(t), u_I(t)) + \eta_I(x_I(t), \bar{x}_I(t), u_I(t), t) \\ \quad + \beta_I^x(t - T_0^x) \phi_I(x(t), u_I(t)) & (1) \\ y_I(t) = C_I x_I(t) + \xi_I(t) + \beta_I^y(t - T_0^y) \theta_I(t) & (2) \end{cases}$$

where $t \in \mathbb{N}$ is the discrete time instant, $x_I \in \mathbb{R}^{n_I}$, $u_I \in \mathbb{R}^{m_I}$ and $y_I \in \mathbb{R}^{p_I}$ are the state, input and measured output vectors of the I -th subsystem, respectively, and $x \triangleq [x_1^\top, x_2^\top, \dots, x_N^\top]^\top \in \mathbb{R}^n$ is the state vector of the overall system. Note that the distributed fault diagnosis scheme (to be presented) is composed by N local fault detection agents \mathcal{F}_I , $I \in \{1, \dots, N\}$, one for each subsystem Σ_I , and, that the structure of the diagnosis agents is exactly mirroring the decomposition (1), (2). The matrix $A_I \in \mathbb{R}^{n_I} \times \mathbb{R}^{n_I}$ and the function $g_I : \mathbb{R}^{p_I} \times \mathbb{R}^{p_I} \times \mathbb{R}^{m_I} \mapsto \mathbb{R}^{n_I}$ are the known nominal

function dynamics and the matrix $C_I \in \mathbb{R}^{p_I} \times \mathbb{R}^{n_I}$ is the known nominal output matrix of the I -th subsystem. The function g_I contains the known part of the interconnection function between the I -th and its neighboring subsystems. More specifically, the vectors $y_I^0(t)$, $\bar{y}_I^0(t)$ are defined as $y_I^0(t) \triangleq C_I x_I(t)$ and $\bar{y}_I^0(t) \triangleq \bar{C}_I \bar{x}_I(t)$, where $\bar{x}_I \in \mathbb{R}^{\bar{n}_I}$ and $\bar{C}_I \bar{x}_I \in \mathbb{R}^{\bar{p}_I}$ denote the state variables and the corresponding output variables, respectively, of the neighboring subsystems that affect the I -th subsystem. This indicates that g_I is a function of local variables $y_I^0(t)$ and, interconnection variables $\bar{y}_I^0(t)$ that are measurable as $y_I(t)$, $\bar{y}_I(t)$ respectively. The superscript 0 in $y_I^0(t)$, $\bar{y}_I^0(t)$ simply indicates the noiseless and sensor fault free measurements of $y_I(t)$, $\bar{y}_I(t)$ respectively. The vector function $\eta_I : \mathbb{R}^{n_I} \times \mathbb{R}^{\bar{n}_I} \times \mathbb{R}^{m_I} \times \mathbb{R}^+ \mapsto \mathbb{R}^{n_I}$ denotes the modeling uncertainty associated with the nominal dynamics and $\xi_I \in \mathbb{R}^{p_I}$ represents the measurement noise. The term $\beta_I^x(t - T_0^x) \phi_I(x, u_I)$ characterizes the process fault function dynamics affecting the I -th subsystem, including its time evolution. More specifically, the term $\phi_I : \mathbb{R}^n \times \mathbb{R}^{m_I} \mapsto \mathbb{R}^{n_I}$ represents the unknown fault function and the term $\beta_I^x(t - T_0^x) : \mathbb{R} \mapsto \mathbb{R}^+$ models the time evolution of the fault, where T_0^x is the unknown time of the fault occurrence. Note that the fault function ϕ_I may depend on the global state variable vector x and not only on the local state vector x_I allowing faults to be functions of the overall state vector and not only of the states that are available to the I -th subsystem. The term $\beta_I^y(t - T_0^y) \theta_I(t)$ characterizes the sensor fault, where $\beta_I^y : \mathbb{R} \mapsto \mathbb{R}^+$ models the time profile of the sensor fault which occurs at some unknown time T_0^y and, $\theta_I \in \mathbb{R}^{p_I}$ represents the unknown time-varying sensor fault. In this work, no particular modeling is considered for the time profile $\beta_I^x(t - T_0^x)$ of the process fault and $\beta_I^y(t - T_0^y)$ of the sensor fault. Generally, the time profiles can be used to model both abrupt and incipient faults. In this work we consider them to be zero prior to the respective fault occurrence and do not make any modeling considerations regarding the fault evolution after their occurrence, i.e. we only consider $\beta_I^x(t - T_0^x) = 0 \quad \forall t < T_0^x$ and $\beta_I^y(t - T_0^y) = 0 \quad \forall t < T_0^y$. In fact, the faults can be permanent, temporary or even intermittent.

In this work, subsystem Σ_J is said to affect subsystem Σ_I (or in other words Σ_J is a neighbor of Σ_I), if the interconnection variables of Σ_I , i.e. $\bar{y}_I^0(t)$, contains at least one of the measurable output variables of Σ_J , i.e. $y_J^0(t)$.

The objective is to design and analyze a distributed fault detection approach where, a local fault detection agent \mathcal{F}_I is associated with each subsystem Σ_I and receives local measurements u_I , y_I and partial information from neighboring fault detection agents \mathcal{F}_J . Each fault detection agent \mathcal{F}_I is not connected to all other agents, but only to a subset of neighboring agents, thus constituting a distributed fault detection scheme. The analysis part does not aim only in the derivation of suitable detection thresholds but also, it aims in the investigation of the fault propagation issues. In this work, the notion of fault propagation does not mean the creation of new faults in interconnected subsystems as a result of a faulty behavior of a subsystem. Instead, it means the way a particular fault occurring in one subsystem affects neighboring interconnected subsystems (in other words, it is simply the propagation of the fault effects from a faulty subsystem to its interconnected subsystems it affects). More specifically, the objective is to design a robust distributed fault detection scheme, for process and sensor faults, with enhanced fault detectability characteristics and, inherent fault isolation characteristics, that is able to provide information regarding the type and location of the fault that has occurred. The fault detectability enhancement is achieved through filtering by exploiting the noise suppression properties of filters in order to obtain tight detection thresholds. The enhanced fault isolation characteristics are integrated by the design of the scheme which utilizes the measurements instead of the state estimates. The purpose is to derive a high-level isolation scheme, which does not necessarily pinpoints the specific fault that has occurred, but rather infers some conclusions about the fault that has occurred by taking into consideration the decisions of the fault detection agents that monitor each subsystem. This information, which in some special cases could even lead to the identification of a faulty sensor, can provide valuable information that can be used by a more advanced fault isolation scheme in order to greatly improve its performance by excluding potential fault scenarios. In the sequel, the terms fault isolation/diagnosis will be used according to the aforementioned basis.

Each fault detection agent contains an estimation model based on its subsystems' nominal dy-

namics that provides the state estimates and, utilizes filtering to derive the residual and threshold signals. Finally, each detection agent provides a binary decision regarding the detection of a fault in the subsystem it monitors. The decisions of all the fault detection agents are then exploited by the high-level isolation scheme to infer some information regarding the type/location of the fault that has occurred.

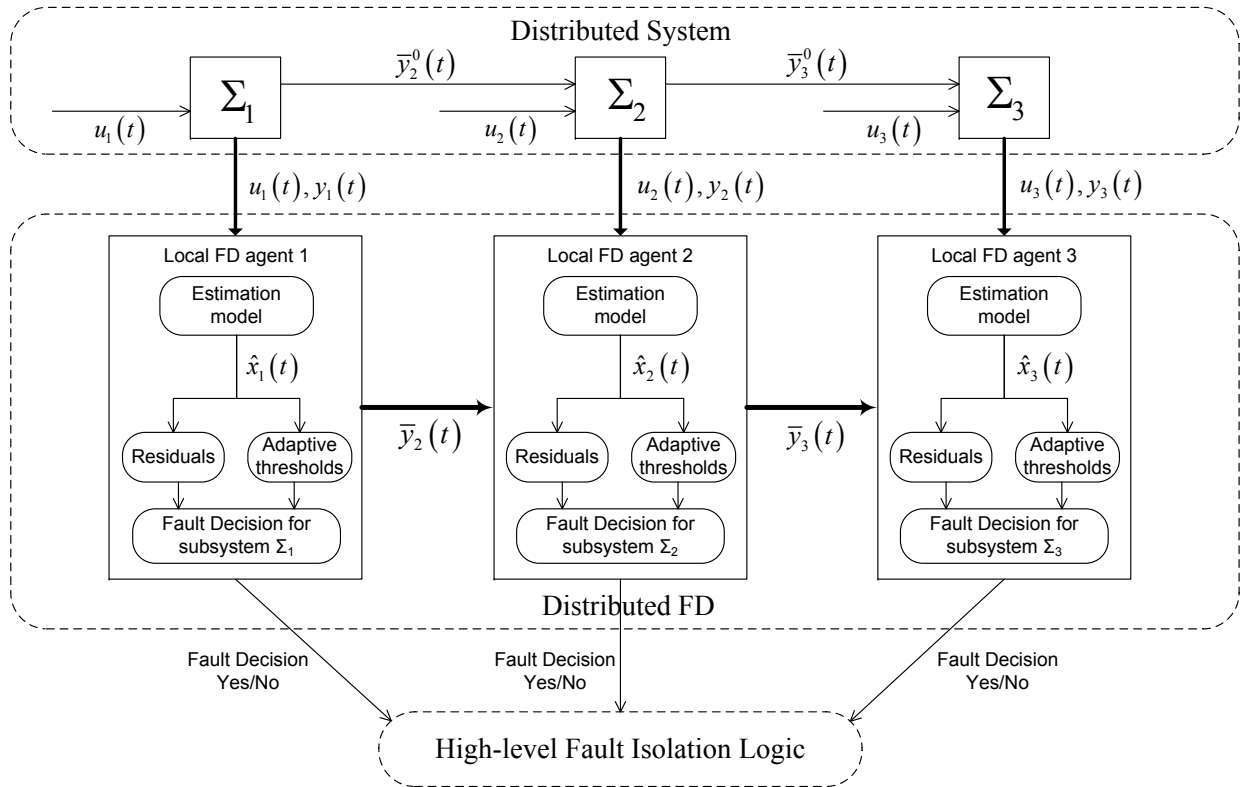


Figure 1.: Distributed fault diagnosis approach for the case of three subsystems Σ_1 , Σ_2 , Σ_3 where Σ_1 affects Σ_2 and Σ_2 affects Σ_3 .

Generally, the distributed fault diagnosis scheme is composed by N local fault detection agents \mathcal{F}_I , $I = 1, \dots, N$, one for each subsystem Σ_I . Each local fault detection agent \mathcal{F}_I requires the input and output measurements of the subsystem Σ_I that is monitoring and also the measurements of all interconnecting subsystems Σ_J that are affecting Σ_I . Note that these last measurements are communicated by neighboring fault detection agents \mathcal{F}_J , and not by the subsystems Σ_J . Therefore, there is the need of communication between the fault detection agents depending on their interconnections which constitutes the scheme distributed. Note that, the information exchanged among the subsystems is constituted only by quantities $(\bar{y}_I^0(t))$ that are measurable with some uncertainty $(\bar{y}_I(t))$. Figure 1 illustrates the distributed fault diagnosis approach for the case of three subsystems Σ_1 , Σ_2 , Σ_3 where Σ_1 affects Σ_2 and Σ_2 affects Σ_3 .

The fault diagnosis structure can also be considered as a hierarchical multi-agent diagnostic system composed of two layers: a lower and an upper layer. In the lower layer, a diagnostic agent \mathcal{F}_I is associated to each subsystem Σ_I with the aim to detect process and sensor faults. Each fault detection agent \mathcal{F}_I consists of a residual generator of the form (4), (5), (6) (to be given in the sequel), together with a fault detection decision logic relying on the comparison of the residual (6) to an adaptive threshold (9) (to be designed). In the upper layer, a fault isolation logic combines the decisions of the diagnostic agents, with the aim to infer further information regarding the type of the fault that has occurred (distinguish process and sensor faults) and its location.

In this work, we do not deal explicitly with the design of feedback controllers for selecting u_I . Instead, we consider the fault detection issue in the presence of process faults ϕ_I , sensor faults

θ_I , modeling uncertainties η_I and measurement noise ξ_I . The proposed formulation allows for any controllers that achieve under healthy conditions some desired control objectives and does not depend on their structure. It is assumed that the controllers are able to retain the uniform boundedness of the state variables before and after the occurrence of a fault.

The following assumptions are used throughout the paper:

Assumption 1: For each subsystem Σ_I , $I \in \{1, \dots, N\}$, the local state variables $x_I(t)$ and the local inputs $u_I(t)$ belong to a known compact region \mathcal{D}_{x_I} and \mathcal{D}_{u_I} respectively before and after the occurrence of a fault, i.e. $x_I(t) \in \mathcal{D}_{x_I}$, $u_I(t) \in \mathcal{D}_{u_I}$ for all $t \geq 0$.

Assumption 2: For each subsystem Σ_I , $I \in \{1, \dots, N\}$ the pair (A_I, C_I) is detectable.

Assumption 3: The modeling uncertainty η_I in each subsystem is an unstructured and unknown nonlinear function of x_I , \bar{x}_I , u_I and t but bounded by a known positive functional $\bar{\eta}_I$ under sensor fault-free operation, i.e.,

$$\|\eta_I(x_I, \bar{x}_I, u_I, t)\| \leq \bar{\eta}_I(y_I, \bar{y}_I, u_I), \quad (3)$$

for all $t \in \mathbb{N}$ and for all $(x_I, \bar{x}_I, u_I) \in \mathcal{D}_I$, where $\bar{y}_I \in \mathbb{R}^{\bar{p}_I}$ is the noisy counterpart of $\bar{y}_I^0(t)$, i.e. $\bar{y}_I = \bar{y}_I^0(t) + \bar{\xi}_I$, $\bar{\xi}_I \in \mathbb{R}^{\bar{p}_I}$ and $\bar{\eta}_I(y_I, \bar{y}_I, u_I) \geq 0$ is a known bounding function in some compact region of interest $\mathcal{D}_I = \mathcal{D}_{x_I} \times \mathcal{D}_{\bar{x}_I} \times \mathcal{D}_{u_I} \subset \mathbb{R}^{n_I} \times \mathbb{R}^{\bar{n}_I} \times \mathbb{R}^{m_I}$.

Assumption 4: The measurement noise belongs to a known compact region, i.e. $\xi_I(t) \in \mathcal{D}_{\xi_I} \subset \mathbb{R}^{p_I}$, $\bar{\xi}_I \in \mathcal{D}_{\bar{\xi}_I} \subset \mathbb{R}^{\bar{p}_I}$.

Assumption 1 is required for well-posedness since in this work we do not address the control design and fault accommodation problem, but instead the fault detection problem. Assumption 2 is required for the design of a suitable observer to be used for the residual signal generation. Assumption 3 characterizes the class of modeling uncertainties being considered. The bound $\bar{\eta}_I$ is required in order to distinguish the effects between modeling uncertainty and faults. Assumption 4 is required in order to distinguish the effects between noise and faults.

3. Distributed Fault Detection

In this Section, the details of the proposed distributed scheme regarding the design of the residual and threshold signals are given, along with some practical considerations.

3.1 Residual Signal Generation

In this part, the residual signal generation in each fault detection agent is addressed by making use of filtering. For each subsystem Σ_I , we consider an estimation model, based on the known components of (1) under healthy mode of operation:

$$\hat{x}_I(t+1) = A_I \hat{x}_I(t) + g_I(y_I(t), \bar{y}_I(t), u_I(t)) + L_I(y_I(t) - \hat{y}_I(t)) \quad (4)$$

$$\hat{y}_I(t) = C_I \hat{x}_I(t), \quad (5)$$

where the gain matrix L_I is computed so that $(A_I - L_I C_I)$ is Schur stable, i.e. its eigenvalues lie in the open unit disc. Note that, since the pair (A_I, C_I) is detectable, according to Assumption 2, such L_I can always be determined. In order to simplify the presentation of the mathematical calculations, the initial condition $\hat{x}_I(0)$ is considered known as $\hat{x}_I(0) = x_I(0)$. In the case $x_I(0)$ is not exactly known, the discrepancy $x_I(0) - \hat{x}_I(0)$ will appear in the calculations which, as it will be shown later,

is multiplied by exponentially decaying functions, and therefore it does not affect substantially the subsequent analysis.

In this work, the residual signal $r_I(t)$ to be used for fault detection in each subsystem Σ_I is given by

$$r_I(t) \triangleq H(z) [y_I(t) - \hat{y}_I(t)], \quad (6)$$

where $H(z)$ is a p -th order, asymptotically stable filter with proper transfer function

$$H(z) = \frac{d_0 + d_1 z^{-1} + d_2 z^{-2} + \dots + d_p z^{-p}}{1 + c_1 z^{-1} + \dots + c_p z^{-p}}. \quad (7)$$

Note that the form of $H(z)$ allows both IIR and FIR types of digital filters. In addition, note that, for the residual generation, each measured variable $y_I^{(j)}$ (j -th component of y_I) is filtered by $H(z)$ in order to dampen the effect of measurement uncertainty $\xi_I(t)$, so that tighter detection thresholds can be obtained. Specifically, for each measured variable $y_I^{(j)}$, a corresponding residual $r_I^{(j)}$ and threshold $\bar{r}_I^{(j)}$ are generated and hence, all the measurements need to be filtered for the residual generation. On the other hand, the estimation model given by (4), (5) relies on the unfiltered measurements. In the proposed approach, the same filter $H(z)$ must be used within the fault detection agent \mathcal{F}_I for filtering all the measurements y_I , and different filters $H(z)$ can be used by each fault detection agent. In this work, without loss of generality, we have considered that the same filter $H(z)$ is used by all detection agents.

The choice of a particular type of filter to be used is application dependent, and it is made according to the available a-priori knowledge on the noise properties. Usually, measurement noise is constituted by high frequency components and therefore the use of low-pass filter for dampening noise is well justified. On other occasions, one may want to focus the fault detectability on a prescribed frequency band of the measurement signals and hence, choose the filter accordingly. The particular selection criteria for choosing a suitable filter and its trade-offs are out of the scope of the present paper and the reader is referred to the continuous-time case in Keliris et al. (2013a) where a rigorous investigation of the filtering impact (according to the poles' location and filters' order) on the detection time is presented.

Since the filter $H(z)$ is asymptotically stable, for bounded measurement noise $\xi_I(t)$, the filtered measurement noise $\epsilon_{\xi_I}(t) \triangleq H(z) [\xi_I(t)]$ is bounded as follows:

$$\|\epsilon_{\xi_I}(t)\| \leq \bar{\epsilon}_{\xi_I}(t), \quad (8)$$

where $\bar{\epsilon}_{\xi_I}$ is a computable bounding function. Depending on the noise characteristics, $H(z)$ can be selected to reduce the bounding function $\bar{\epsilon}_{\xi_I}$. It is important to note that filtering is primarily used to mitigate the effect of measurement noise and aid in the derivation of tighter thresholds, thus enhancing fault detectability (see Keliris et al. (2013a)).

Remark 1: It is important to note that in the nonlinear function g_I , the measurements y_I and \bar{y}_I are used instead of the estimates $C_I \hat{x}_I$ and $\bar{C}_I \hat{\hat{x}}_I$. This is crucial for the derivation of the high-level fault isolation scheme, since in the case of a process fault its effects are contained in the measurements, and as it will be shown, the fault can only be detected by the agent monitoring the subsystem that the fault has occurred. On the other hand, when a sensor fault occurs, it can affect neighboring detection agents through the communicated measurements of the interconnection variables, i.e. \bar{y}_I that contain the sensor fault and hence the fault can also be detected by interconnected detection agents. In X. Zhang et al. (2009), which investigated a decentralized fault detection scheme for process faults, the estimation model used the estimates in the interconnection functions among subsystems (instead of the measurements as in this work) in order to enhance fault detectability by allowing the interconnected agents to also be able to detect the fault. In this work, the measurements are used instead of the estimates, in order to enhance fault isolability.

3.2 Adaptive Detection Threshold

In this part, suitable detection thresholds that guarantee no false alarms are derived. Filtering is also integrated in the design in order to attenuate the measurement noise effects and aid in the derivation of tighter detection thresholds. This is achieved by treating the use of filter as a linear state transformation which leads to the manipulation of the mathematical expressions based on their filtered versions.

In the following, we will use the filtered state variable $x_{I,f}(t) \triangleq H(z)[x_I(t)]$, and the filtered state estimate $\hat{x}_{I,f}(t) \triangleq H(z)[\hat{x}_I(t)]$. The threshold design is based on the derivation of a suitable bound on the filtered state estimation error $x_{I,f}(t) - \hat{x}_{I,f}(t)$. Also, let $h(t)$ be the impulse response associated with $H(z)$; i.e. $h(t) \triangleq \mathcal{Z}^{-1}[H(z)]$, so that $x_{I,f}(t)$ can be written as

$$x_{I,f}(t) = \sum_{k=0}^t h(k)x_I(t-k).$$

In this work, the detection decision of a fault in the overall system is made when $|r_I^{(j)}(t)| > \bar{r}_I^{(j)}(t)$ at some time t for at least one component $j \in \{1, 2, \dots, p_I\}$ in any local subsystem Σ_I , where $\bar{r}_I^{(j)}(t)$ is the detection threshold given by

$$\bar{r}_I^{(j)}(t) \triangleq \sum_{k=0}^{t-1} \alpha_{I,j} \delta_{I,j}^{t-1-k} \bar{\chi}_I(k) + \bar{\epsilon}_{\xi_I}(t), \quad (9)$$

where

$$\bar{\chi}_I(t) \triangleq \bar{H}(z) \left[\bar{\eta}_I(y_I(t), \bar{y}_I(t), u_I(t)) + \bar{\Delta}g_I \right] + \|L_I\| \bar{\epsilon}_{\xi_I}(t), \quad (10)$$

$\bar{H}(z)$ is a filter with impulse response $\bar{h}(t) \geq |h(t)|$ for all $t \geq 0$,

$$\bar{\Delta}g_I \triangleq \sup_{\substack{(x_I, \bar{x}_I, u_I) \in \mathcal{D}_I \\ (\xi_I, \bar{\xi}_I) \in \mathcal{D}_{\xi_I} \times \mathcal{D}_{\bar{\xi}_I}}} \|g_I(C_I x_I, \bar{C}_I \bar{x}_I, u_I) - g_I(C_I x_I + \xi_I, \bar{C}_I \bar{x}_I + \bar{\xi}_I, u_I)\|, \quad (11)$$

and where the constants $\alpha_{I,j} > 0$ and $0 < \delta_{I,j} < 1$ are selected so that the following inequality holds:

$$\|C_I^{(j)} A_{I,0}^t\| \leq \alpha_{I,j} \delta_{I,j}^t \leq \|C_I^{(j)}\| \|A_{I,0}\|^t, \quad (12)$$

with $A_{I,0} \triangleq A_I - L_I C_I$. Finally, $C_I^{(j)}$ denotes the j -th row of the matrix C_I . Note that, since $A_{I,0}$ is Schur stable, suitable constants $\alpha_{I,j}$, $\delta_{I,j}$ do exist (Ferrari, Parisini, & Polycarpou, 2008).

Note that, the threshold (9) can be implemented using linear filtering techniques as:

$$\bar{r}_I^{(j)}(t) = \frac{\alpha_{I,j}}{1 - \delta_{I,j} z^{-1}} [\bar{\chi}_I(t-1)] + \bar{\epsilon}_{\xi_I}(t). \quad (13)$$

Figure 2 illustrates the implementation of the fault detection scheme for the detection agent \mathcal{F}_I resulting from equations (7), (4), (5), (6) and (13).

In the absence of any faults, the residual signal $r_I^{(j)}(t)$ given in (6) is always bounded by the detection threshold $\bar{r}_I^{(j)}(t)$ given by (9). The fault detection concept is formalized in the following Lemma.

Lemma 1: Consider a distributed system made of N subsystems Σ_I given by (1), (2). In the absence of any faults, the residuals $r_I^{(j)}(t)$ given by (6), where $\hat{y}_I^{(j)}$ are given by (4) and (5), are bounded by

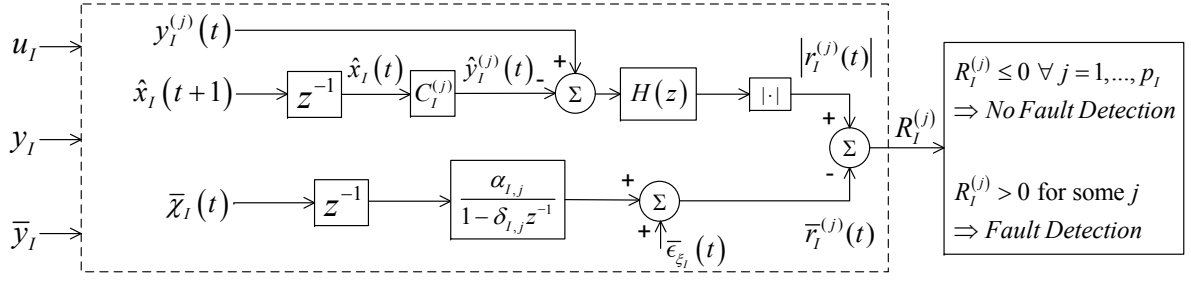


Figure 2.: Fault detection scheme of the detection agent \mathcal{F}_I .

the detection thresholds $\bar{r}_I^{(j)}(t)$, given by (9), thus guaranteeing that no false alarms are issued by the fault detection scheme.

Proof. The filtered state estimate $x_{I,f}(t+1)$ can be written as:

$$\begin{aligned} x_{I,f}(t+1) &= \sum_{k=0}^{t+1} h(t+1-k)x_I(k) \\ &= \sum_{k=1}^{t+1} h(t+1-k)x_I(k) + h(t+1)x_I(0), \end{aligned}$$

and, by using the change of variables $k = i + 1$, it becomes

$$x_{I,f}(t+1) = \sum_{i=0}^t h(t-i)x_I(i+1) + h(t+1)x_I(0). \quad (14)$$

i) Prior to the possible occurrence of a process fault, by using (1), (14) can be written as:

$$\begin{aligned} x_{I,f}(t+1) &= \sum_{i=0}^t h(t-i)[A_I x_I(i) + g_I(y_I^0(i), \bar{y}_I^0(i), u_I(i)) + \eta_I(x_I(i), \bar{x}_I(i), u_I(i), i)] + h(t+1)x_I(0) \\ &= A_I x_{I,f}(t) + H(z) [g_I(y_I^0(t), \bar{y}_I^0(t), u_I(t))] + H(z) [\eta_I(x_I(t), \bar{x}_I(t), u_I(t), t)] \\ &\quad + h(t+1)x_I(0). \end{aligned} \quad (15)$$

Similarly to the derivation of (15), the filtered state estimate dynamics $\hat{x}_{I,f}(t+1)$ by using (4) satisfies:

$$\begin{aligned} \hat{x}_{I,f}(t+1) &= A_I \hat{x}_{I,f}(t) + H(z) [g_I(y_I(t), \bar{y}_I(t), u_I(t))] + L_I (y_{I,f}(t) - \hat{y}_{I,f}(t)) \\ &\quad + h(t+1)\hat{x}_I(0), \end{aligned} \quad (16)$$

where $y_{I,f}(t) \triangleq H(z)[y_I(t)]$ and $\hat{y}_{I,f}(t) \triangleq H(z)[\hat{y}_I(t)] = C_I \hat{x}_{I,f}(t)$.

ii) Prior to the possible occurrence of a sensor fault, $y_{I,f}(t) = C_I x_{I,f}(t) + \epsilon_{\epsilon_I}(t)$ and, by using (15), (16) and some algebra, the filtered state estimation error $\tilde{x}_{I,f}(t) \triangleq x_{I,f}(t) - \hat{x}_{I,f}(t)$ satisfies

$$\tilde{x}_{I,f}(t+1) = A_{I,0} \tilde{x}_{I,f}(t) + \chi_I(t), \quad (17)$$

where

$$\chi_I(t) \triangleq H(z) [\eta_I(x_I(t), \bar{x}_I(t), u_I(t), t) + \Delta g_I(t)] - L_I \epsilon_{\xi_I}(t) + h(t+1)(x_I(0) - \hat{x}_I(0)), \quad (18)$$

$$\Delta g_I(t) \triangleq g_I(y_I^0(t), \bar{y}_I^0(t), u_I(t)) - g_I(y_I^0(t) + \xi_I(t), \bar{y}_I^0(t) + \bar{\xi}_I(t), u_I(t)). \quad (19)$$

The solution of (17) is

$$\tilde{x}_{I,f}(t) = A_{I,0}^t \tilde{x}_{I,f}(0) + \sum_{k=0}^{t-1} A_{I,0}^{t-1-k} \chi_I(k). \quad (20)$$

Note that in (18), $h(t+1)(x_I(0) - \hat{x}_I(0)) = 0$ because of $x_I(0) = \hat{x}_I(0)$, and note that the term $\tilde{x}_{I,f}(0)$ in (20) is also zero since $\tilde{x}_{I,f}(0) = x_{I,f}(0) - \hat{x}_{I,f}(0) = h(0)(x_I(0) - \hat{x}_I(0)) = 0$. In the case $\hat{x}_I(0) \neq x_I(0)$, the aforementioned terms decay exponentially to zero and hence do not affect substantially the subsequent analysis. More specifically, the term $A_{I,0}^t \tilde{x}_{I,f}(0)$ in (20) decays exponentially to zero because $A_{I,0}$ is Schur stable, and the term $h(t+1)(x_I(0) - \hat{x}_I(0))$ in (18) decays also exponentially to zero because the impulse response $h(t)$ of an asymptotically stable filter is exponentially decaying (in the case of IIR filter, the FIR filter case is trivial).

Now, by using (2) and (5) the residual (6) prior to any fault satisfies $r_I(t) = C_I \tilde{x}_{I,f}(t) + \epsilon_{\xi_I}(t)$ and, by using (20) and $\tilde{x}_{I,f}(0) = 0$, it becomes

$$r_I(t) = \sum_{k=0}^{t-1} C_I A_{I,0}^{t-1-k} \chi_I(k) + \epsilon_{\xi_I}(t). \quad (21)$$

By taking the absolute value component-wise and using the triangle inequality, the j -th element of $r_I(t)$, i.e. $r_I^{(j)}(t)$ satisfies

$$\begin{aligned} |r_I^{(j)}(t)| &\leq \left| \sum_{k=0}^{t-1} C_I^{(j)} A_{I,0}^{t-1-k} \chi_I(k) \right| + |\epsilon_{\xi_I}^{(j)}(t)| \\ &\leq \sum_{k=0}^{t-1} \|C_I^{(j)} A_{I,0}^{t-1-k}\| \|\chi_I(k)\| + |\epsilon_{\xi_I}^{(j)}(t)|. \end{aligned} \quad (22)$$

Moreover, using (12) and the fact that $|\epsilon_{\xi_I}^{(j)}(t)| \leq \|\epsilon_{\xi_I}(t)\| \leq \bar{\epsilon}_{\xi_I}(t)$, (22) becomes

$$|r_I^{(j)}(t)| \leq \sum_{k=0}^{t-1} \alpha_{I,j} \delta_{I,j}^{t-1-k} \|\chi_I(k)\| + \bar{\epsilon}_{\xi_I}(t). \quad (23)$$

Now, consider the term $\chi_I(t)$ which satisfies

$$\begin{aligned} \|\chi_I(t)\| &= \|H(z) [\eta_I(x_I, \bar{x}_I, u_I, t) + \Delta g_I(t)] - L_I \epsilon_{\xi_I}(t)\| \\ &\leq \|H(z) [\eta_I(x_I, \bar{x}_I, u_I, t) + \Delta g_I(t)]\| + \|L_I \epsilon_{\xi_I}(t)\| \\ &\leq \sum_{k=0}^t |h(t-k)| \|\eta_I(x_I(k), \bar{x}_I(k), u_I(k), k)\| + \sum_{k=0}^t |h(t-k)| \|\Delta g_I(k)\| + \|L_I\| \|\epsilon_{\xi_I}(t)\| \\ &\leq \bar{\chi}_I(t) \end{aligned} \quad (24)$$

where $\bar{\chi}_I(t)$ is the bounding function given by (10).

Finally, by using (23), (24) and the bound $\bar{\chi}_I(t)$, we obtain $|r_I^{(j)}(t)| \leq \bar{r}_I^{(j)}(t)$, where the detection threshold $\bar{r}_I^{(j)}(t)$ is given by (9), thus concluding the proof. \square

From a practical viewpoint, the implementation of the threshold \bar{r}_I requires the bound $\bar{\Delta}g_I$ given in (11). One approach to derive this bound is to consider a local Lipschitz condition, i.e.:

$$\|g_I(C_I x_I, \bar{C}_I \bar{x}_I, u_I) - g_I(C_I x_I + \xi_I, \bar{C}_I \bar{x}_I + \bar{\xi}_I, u_I)\| \leq L_{g_I} \|\begin{bmatrix} \xi_I & \bar{\xi}_I \end{bmatrix}^\top\|,$$

where L_{g_I} is the Lipschitz constant for the function g_I with respect to (x_I, \bar{x}_I) in the region $\mathcal{D}_{x_I} \times \mathcal{D}_{\bar{x}_I}$. Therefore, by using a uniform bound on the measurement noise (see Assumption 4), i.e. $\|\xi_I(t)\| \leq \xi_{I,b}$, $\|\bar{\xi}_I(t)\| \leq \bar{\xi}_{I,b} \quad \forall t \in \mathbb{N}$, then we can derive the bound $\bar{\Delta}g_I$.

Filtering is primarily used to dampen the measurement noise and allow the derivation of tighter detection thresholds. Filtering can also be proved beneficial for dampening the mismatch function $\Delta g_I(t)$ which results due to the measurement noise and therefore further enhance fault detectability. Among the various filters $H(z)$ one can select, some may lead to less conservative detection thresholds. The derivation of potentially tighter thresholds is obtained by writing the total uncertainty term $\chi_I(t)$ given by (18) as

$$\chi_I(t) = H(z) [\eta_I(x_I, \bar{x}_I, u_I, t)] + \epsilon_{\Delta g_I}(t) - L_I \epsilon_{\xi_I}(t), \quad (25)$$

where $\epsilon_{\Delta g_I}(t) \triangleq H(z) [\Delta g_I(t)]$, and by making the following assumption.

Assumption 5: *In the absence of a sensor fault, the filtered function mismatch term $\epsilon_{\Delta g_I}(t)$ is bounded by a computable positive function $\bar{\epsilon}_{\Delta g_I}(t)$; i.e., for all $t \in \mathbb{N}$,*

$$\|\epsilon_{\Delta g_I}(t)\| \leq \bar{\epsilon}_{\Delta g_I}(t). \quad (26)$$

Assumption 5 is based on the fact that filtering dampens the error effect of measurement noise present in the function mismatch term $\Delta g_I(t)$. A suitable selection of $\bar{\epsilon}_{\Delta g_I}$ can be made through the use of simulations (i.e. Monte Carlo methods) by filtering the function mismatch term $\Delta g_I(t)$ using the known nominal function dynamics and the available noise characteristics (recall that the measurement noise is assumed to take values in a compact set, see Assumption 4).

In this case, the detection threshold $\bar{r}_I^{(j)}(t)$ is still given by (9), but $\bar{\chi}_I(t)$ is now given by:

$$\bar{\chi}_I(t) = \bar{H}(z) [\bar{\eta}_I(y_I(t), \bar{y}_I(t), u_I(t))] + \bar{\epsilon}_{\Delta g_I}(t) + \|L_I\| \bar{\epsilon}_{\xi_I}(t). \quad (27)$$

As a result, the detection threshold by using (27) can be less conservative than by using (10). More information regarding this can be found in Keliris et al. (2013a).

3.3 Selection of filter $\bar{H}(z)$

In this part, we give more details on the selection of a suitable filter $\bar{H}(z)$ which is required for the implementation of the detection threshold. As stated before, its impulse response must satisfy $|h(t)| \leq \bar{h}(t)$ for all $t \geq 0$. In the case where the impulse response $h(t)$ is non-negative, the selection $\bar{H}(z) = H(z)$ is trivial. Sufficient conditions for non-negative impulse response for a class of discrete-time transfer functions are given in Liu and Bauer (2008). In the following, we briefly illustrate two simple methods for choosing $\bar{H}(z)$, one considering $H(z)$ as a digital IIR filter and the other one as a FIR filter.

First we consider the case where $H(z)$ is an IIR filter. As stated earlier, the impulse response $h(t)$ of a proper and asymptotically stable transfer function $H(z)$ converges to zero exponentially fast. Therefore, there exist $\kappa > 0$, $\lambda \in [0, 1)$ such that for all $t \in \mathbb{N}$ the following inequality holds:

$|h(t)| \leq \kappa\lambda^t$. Since $|h(t)| \leq \bar{h}(t)$ must hold, the impulse response $\bar{h}(t)$ can be selected as $\bar{h}(t) = \kappa\lambda^t$ and thus $\bar{H}(z) = \frac{\kappa}{1-\lambda z^{-1}}$.

Now, let's consider the case in which $H(z)$ is a FIR filter. Let $H(z)$ be a p -th order FIR filter given by $H(z) = \sum_{k=0}^p d_k z^{-k}$. Therefore, $\bar{h}(t)$ can be selected as $\bar{h}(t) = |h(t)|$ which leads to the FIR filter $\bar{H}(z) = \sum_{k=0}^p |d_k| z^{-k}$.

Generally, in fault detection schemes it is difficult to find a suitable good balance in selecting the threshold. If the threshold is too high (conservative threshold) then some faults may go undetected. If the threshold is too low, then this may result in false alarms. The following Remark discusses the sources of possible conservativeness in the proposed scheme.

Remark 2: The scheme presented in this paper, guarantees no-false alarms but, at the same time, utilizes filtering to dampen the measurement noise and hence, improve the derivation of tighter detection thresholds for enhanced fault detectability. Specifically, the designed fault detection scheme guarantees that in the absence of a fault, the residual $r_I(t)$ given by (6) is uniformly bounded by the detection threshold $\bar{r}_I(t)$ given by (9), which requires the bound $\bar{\chi}_I(t)$. The main sources of conservativeness in the designed threshold are included in the overall bound $\bar{\chi}_I(t)$, which can be broken up into the following four components: a) the filter $\bar{H}(z)$, b) the bound on the modeling uncertainty $\eta_I(x_I, \bar{x}_I, u_I)$, c) the bound on the filtered noise $\epsilon_{\xi_I}(t)$ and d) the bound on the function mismatch term $\Delta g_I(t)$. At first, the filter $\bar{H}(z)$, which is required to satisfy $|h(t)| \leq \bar{h}(t)$ for all $t \geq 0$, may impose some conservativeness but, this can be avoided by selecting the filter $H(z)$ used for filtering the measurements to have a non-negative impulse response so that the same filter can be used for the threshold derivation, i.e. $\bar{H}(z) = H(z)$. In this case, no conservativeness in the threshold is added. The second source of conservativeness stems from the bound on $\eta_I(x_I, \bar{x}_I, u_I)$, which according to Assumption 3, is bounded by a known function $\bar{\eta}_I(y_I, \bar{y}_I, u_I)$. This is required in order to distinguish the effects from modeling uncertainty and faults so that no false alarms are introduced. In practice, the system can be modeled more accurately in certain regions of the state space and therefore, the fact that the bound $\bar{\eta}_I$ is a function of y_I, \bar{y}_I and u_I provides more flexibility by allowing the designer to take into consideration any prior knowledge of the system. Regarding the conservativeness imposed by the use of the bound on the filtered noise $\epsilon_{\xi_I}(t)$, please note that the bound $\bar{\epsilon}_{\xi_I}(t)$ is multiplied with $\|L_I\|$ in (10) and hence, the conservativeness is significantly reduced in comparison to the case in which no filtering is used (in that case, the bound $\bar{\xi}_I(t)$ on the noise (i.e. $\|\xi_I(t)\| \leq \bar{\xi}_I(t)$) would be multiplied with $\|L_I\|$, leading to more conservative thresholds). Finally, the last source of conservativeness in the threshold is introduced by the bound on the mismatch function Δg_I given by (19) for which the bound $\bar{\Delta}g_I$ given by (11) is required. As discussed earlier, one way to derive this bound is through the use of the Lipschitz assumption. The use of filtering can aid in the elimination of (some of) the conservativeness imposed by the bound $\bar{\Delta}g_I$, by exploiting the noise suppression properties through Assumption 5 and by using the bound $\bar{\chi}_I(t)$ given by (27).

In Section 4, fault detectability conditions for the aforementioned fault detection scheme are addressed. The conditions given in Section 4 refer to the case of a fault occurring in subsystem Σ_I and being detected by its respective local fault detection agent \mathcal{F}_I . In Section 5, fault propagation from one subsystem to the other is investigated by examining the way the fault effects appear and affect neighboring interconnected subsystems.

4. Local fault detectability analysis

The fault detectability analysis constitutes a theoretical result that characterizes quantitatively (and in implicit way) the class of faults detectable by the proposed scheme. In order to derive the fault detectability conditions, we take into consideration the distinct occurrence of a process fault ϕ_I at an unknown time $t = T_0^x$ or the occurrence of a sensor fault θ_I at an unknown time $t = T_0^y$.

Theorem 1 (Local Process Fault Detectability): *Consider the nonlinear interconnected system (1), (2) with the distributed fault detection scheme described in (4), (5), (6), (9) and (10). A process fault in the I -th subsystem occurring at $t = T_0^x$ is detectable by the respective local fault detection agent \mathcal{F}_I if the filtered process fault function $\phi_{I,f}(x(t), u_I(t), t) \triangleq H(z)[\beta_I^x(t - T_0^x)\phi_I(x(t), u_I(t))]$ satisfies the following inequality at some time $t > T_0^x$, for some $j = 1, 2, \dots, p_I$:*

$$\left| \sum_{k=T_0^x}^{t-1} C_I^{(j)} A_{I,0}^{t-1-k} \phi_{I,f}(x(k), u_I(k), k) \right| > 2\bar{r}_I^{(j)}(t). \quad (28)$$

Proof. In the presence of a process fault that occurs in the I -th subsystem at $t = T_0^x$, (20) becomes:

$$\tilde{x}_{I,f}(t) = A_{I,0}^t \tilde{x}_{I,f}(0) + \sum_{k=0}^{t-1} A_{I,0}^{t-1-k} (\chi_I(k) + \phi_{I,f}(x(k), u_I(k), k)),$$

and by using $\tilde{x}_{I,f}(0) = 0$, the residual $r_I(t) = C_I \tilde{x}_{I,f}(t) + \epsilon_{\xi_I}(t)$ becomes (similarly to (21))

$$r_I(t) = \sum_{k=0}^{t-1} C_I A_{I,0}^{t-1-k} (\chi_I(k) + \phi_{I,f}(x(k), u_I(k), k)) + \epsilon_{\xi_I}(t).$$

By using the triangle inequality, the j -th element of $r_I(t)$ for $t > T_0^x$ satisfies:

$$|r_I^{(j)}(t)| \geq - \left| \sum_{k=0}^{t-1} C_I^{(j)} A_{I,0}^{t-1-k} \chi_I(k) \right| - |\epsilon_{\xi_I}^{(j)}(t)| + \left| \sum_{k=T_0^x}^{t-1} C_I^{(j)} A_{I,0}^{t-1-k} \phi_{I,f}(x(k), u_I(k), k) \right|. \quad (29)$$

Following a similar procedure as in the derivation of the detection threshold (9), (29) becomes

$$|r_I^{(j)}(t)| \geq -\bar{r}_I^{(j)}(t) + \left| \sum_{k=T_0^x}^{t-1} C_I^{(j)} A_{I,0}^{t-1-k} \phi_{I,f}(x(k), u_I(k), k) \right|.$$

For fault detection, the inequality $|r_I^{(j)}(t)| > \bar{r}_I^{(j)}(t)$ must hold for some $j = 1, 2, \dots, p_I$, so the final fault detectability condition given in (28) is obtained. \square

Theorem 2 (Local Sensor Fault Detectability): *Consider the nonlinear interconnected system (1), (2) with the distributed fault detection scheme described in (4), (5), (6), (10) and (9). A sensor fault in the I -th subsystem occurring at $t = T_0^y$ is detectable by the respective local fault detection agent \mathcal{F}_I if the filtered sensor fault function*

$$\theta_{I,f}(t) \triangleq H(z)[\beta_I^y(t - T_0^y)\theta_I(t)],$$

and the following mismatch function $\Delta g_I'$ due to the sensor fault

$$\Delta g_I'(t) \triangleq g_I(y_I^0(t), \bar{y}_I^0(t), u_I(t)) - g_I(y_I^0(t) + \xi_I(t) + \beta_I^y(t - T_0^y)\theta_I(t), \bar{y}_I^0(t) + \bar{\xi}_I(t), u_I(t)),$$

satisfy the following inequality at some time $t > T_0^y$, for some $j = 1, 2, \dots, p_I$:

$$\left| \theta_{I,f}^{(j)}(t) + \sum_{k=0}^{t-1} C_I^{(j)} A_{I,0}^{t-1-k} (H(z)[\Delta g_I'(k)] - L_I \theta_{I,f}(k)) \right| > \bar{r}_I^{(j)}(t) + \bar{\epsilon}_{\xi_I}(t) + \sum_{k=0}^{t-1} \alpha_{I,j} \delta_{I,j}^{t-1-k} (\|H(z)[\eta_I(x_I, \bar{x}_I, u_I, k)]\| + \|L_I\| \bar{\epsilon}_{\xi_I}(k)). \quad (30)$$

Proof. See Appendix A. □

Theorems 1 and 2 provide sufficient conditions for the implicit characterization of certain classes of faults that can be detected by the proposed fault detection scheme. Clearly, the fault functions $\phi_I(x, u_I)$ and θ_I are typically unknown and therefore these conditions cannot be checked a-priori.

Remark 3: The use of filtering is of crucial importance in order to derive tight detection thresholds that guarantee no false alarms (see Keliris et al. (2013a)). As it can be seen in the detectability conditions given by (28), (30) the detection of the fault depends on the filtered process fault function ϕ_I and filtered sensor fault θ_I and as a result, the selection of the filter is very important. Therefore, some filter selections may lead to less conservative thresholds than others.

5. Fault propagation

In this section, fault propagation from one subsystem to the other is investigated and further intuition regarding the isolation properties of the proposed fault diagnosis scheme is obtained. The notion of fault propagation does not mean the creation of additional faults to neighboring subsystems as a result of a faulty behavior in one of them. Instead, it means the way the fault effects in one subsystem appear and affect its neighboring subsystems. More specifically, we consider a fault that occurs in subsystem Σ_J which affects Σ_I and investigate the possibility of fault detection, not by the local fault detection agent \mathcal{F}_J (which may obviously detect the fault), but by the agent \mathcal{F}_I .

The following lemma summarizes the main findings of this analysis.

Lemma 2: *Consider a distributed system made of N subsystems Σ_I given by (1), (2). The distributed fault detection scheme described by the estimation model (4), (5), the residual signals $r_I(t)$ given by (6) and the detection thresholds $\bar{r}_I(t)$ given by (9) guarantees that:*

- (a) *a process fault occurring in subsystem Σ_J which affects Σ_I can only be detected by its corresponding fault detection agent \mathcal{F}_J and not by the detection agent \mathcal{F}_I .*
- (b) *a sensor fault occurring in Σ_J which affects Σ_I can be detected by either the corresponding detection agent \mathcal{F}_J or the detection agent \mathcal{F}_I .*

Proof. The case of the fault (process or sensor) occurring in subsystem Σ_J and being detectable by its corresponding detection agent \mathcal{F}_J has been investigated in Section 3 and respective detectability conditions were given in Section 4. In the sequel, we investigate the possibility of detection of a fault that occurs in Σ_J which affects Σ_I by the detection agent \mathcal{F}_I .

(a) At first, let's consider fault propagation in the case of a process fault. In this case, the process fault effects in Σ_J are propagated to Σ_I through the interconnection variables $\bar{C}_I \bar{x}_I$ (see (1)) and to \mathcal{F}_I through the measurements \bar{y}_I (communicated by \mathcal{F}_J , see (4)). For easier visual indication of the process fault effects that are contained in the interconnection variables \bar{x}_I and the measurements of the output interconnection variables \bar{y}_I , we denote them as $\bar{x}_{I,P}$ and $\bar{y}_{I,P}$ respectively. Note that $\bar{y}_{I,P}(t) = \bar{C}_I \bar{x}_{I,P}(t) + \bar{\xi}_I(t)$. In the case of a process fault occurring in Σ_J which affects Σ_I , the dynamics of Σ_I are given by (1), (2) and the estimation model of the local fault detection agent \mathcal{F}_I by (4), (5). In the aforementioned equations \bar{x}_I and \bar{y}_I are now indicated by $\bar{x}_{I,P}$ and $\bar{y}_{I,P}$ respectively.

Note that, both the dynamics of Σ_I and the estimation model of the local fault detection agent \mathcal{F}_I are affected by the process fault effects that occurred in Σ_J . Following the same analysis as in the proof of Lemma 1, the filtered estimation error still satisfies (17)-(19) (the fault effects enter implicitly through the interconnection variables and their measurements), which are rewritten below with explicit indication of the fault effect through $\bar{x}_{I,P}$ ($\bar{y}_{I,P}$ is expressed in terms of $\bar{x}_{I,P}$):

$$\begin{aligned}\tilde{x}_{I,f}(t+1) &= A_{I,0}\tilde{x}_{I,f}(t) + \chi_{I,P}(t) \\ \chi_{I,P}(t) &\triangleq H(z) [\eta_I(x_I(t), \bar{x}_{I,P}(t), u_I(t), t) + \Delta g_{I,P}(t)] - L_I \epsilon_{\xi_I}(t), \\ \Delta g_{I,P}(t) &\triangleq g_I(C_I x_I(t), \bar{C}_I \bar{x}_{I,P}(t), u_I(t)) - g_I(C_I x_I(t) + \xi_I(t), \bar{C}_I \bar{x}_{I,P}(t) + \bar{\xi}_I(t), u_I(t)).\end{aligned}$$

Now, the residual $r_I(t)$ is ‘‘contaminated’’ with the fault effects that occurred in subsystem Σ_J , and we need to investigate whether this residual is bounded or not by the detection threshold $\bar{r}_I(t)$ given by (9) that is used by the detection agent \mathcal{F}_I . Following the same mathematical calculations as in the derivation of (23), the residual $r_I^{(j)}$ satisfies

$$|r_I^{(j)}(t)| \leq \sum_{k=0}^{t-1} \alpha_{I,j} \delta_{I,j}^{t-1-k} \|\chi_{I,P}(k)\| + \bar{\epsilon}_{\xi_I}(t), \quad (31)$$

where

$$\begin{aligned}\|\chi_{I,P}(t)\| &\leq \bar{H}(z) [\|\eta_I(x_I(t), \bar{x}_{I,P}(t), u_I(t), t)\| + \|\Delta g_{I,P}(t)\|] + \|L_I\| \bar{\epsilon}_{\xi_I}(t) \\ &\leq \bar{H}(z) [\bar{\eta}_I(y_I(t), \bar{y}_{I,P}(t), u_I(t)) + \bar{\Delta} g_I] + \|L_I\| \bar{\epsilon}_{\xi_I}(t).\end{aligned} \quad (32)$$

The last inequality is derived by using the bounds $\|\eta_I(x_I, \bar{x}_{I,P}, u_I, t)\| \leq \bar{\eta}_I(y_I(t), \bar{y}_{I,P}(t), u_I(t))$ (see Assumption 3) and $\|\Delta g_{I,P}(t)\| \leq \bar{\Delta} g_I$ (see Assumption 1 and (11)). Note that, Assumption 3 is stated for sensor fault-free operation because the bound on the modeling uncertainty $\bar{\eta}_I$ makes use of the measurements. In the event of a process fault which changes the state variables, Assumption 3 is still valid since the measurements are essentially these altered state variables (or linear combination, contaminated with the process fault effects) but with some uncertainty due to the measurement noise. In addition, note that the right side of (32) is actually the term $\bar{\chi}_I(t)$ given by (10) which is used by the local fault detection agent \mathcal{F}_I and hence $\|\chi_{I,P}(t)\| \leq \bar{\chi}_I(t)$. Therefore, from (31), it can be seen that the residual still satisfies $|r_I^{(j)}(t)| \leq \bar{r}_I^{(j)}(t)$ for all $j = 1, \dots, p_I$ and hence the fault is not detected by \mathcal{F}_I . In other words, a process fault that occurs in a subsystem can only be detected by its respective fault detection agent.

(b) In this part, we consider fault propagation in the case of a sensor fault that occurs in Σ_J that affects Σ_I . As in the case of a process fault, the local fault detection agent \mathcal{F}_J may detect the fault, but here we investigate the possibility of fault detection by the detection agent \mathcal{F}_I . In this case the sensor faults in Σ_J are propagated to \mathcal{F}_I through the measurements \bar{y}_I (communicated by \mathcal{F}_J , see (4)). For easier visual indication of the measurements \bar{y}_I that contain the sensor faults, we denote them as $\bar{y}_{I,S}$. Therefore, in the case of a sensor fault occurring in Σ_J which affects Σ_I , the dynamics of Σ_I remain unaffected by the sensor fault and are given by (1), (2) whereas the estimation model of the local fault detection agent \mathcal{F}_I is affected by the sensor fault and it is given by (4), (5) where now \bar{y}_I is indicated with $\bar{y}_{I,S}$ (only (4) is affected). Note that $\bar{y}_{I,S} = \bar{y}_I^0(t) + \xi_I(t) + \beta_I^y(t - T_0^y) \theta_I(t)$, where the term $\beta_I^y(t - T_0^y) \theta_I(t)$ indicates the sensor faults that occur in neighboring subsystems and affect Σ_I . Hence, the filtered estimation error still satisfies (17)-(19) (the fault effects enter implicitly through the measurements of the interconnection variables), which are rewritten below with explicit

indication of the fault effect through $\bar{y}_{I,S}$:

$$\begin{aligned}\tilde{x}_{I,f}(t+1) &= A_{I,0}\tilde{x}_{I,f}(t) + \chi_{I,S}(t) \\ \chi_{I,S}(t) &\triangleq H(z) [\eta_I(x_I(t), \bar{x}_I(t), u_I(t), t) + \Delta g_{I,S}(t)] - L_I \epsilon_{\xi_I}(t), \\ \Delta g_{I,S}(t) &\triangleq g_I(y_I^0(t), \bar{y}_I^0(t), u_I(t)) - g_I(y_I^0(t) + \xi_I(t), \bar{y}_{I,S}(t), u_I(t)).\end{aligned}$$

Similarly to the derivation of (23), the residual $r_I^{(j)}$ satisfies

$$|r_I^{(j)}(t)| \leq \sum_{k=0}^{t-1} \alpha_{I,j} \delta_{I,j}^{t-1-k} \|\chi_{I,S}(k)\| + \bar{\epsilon}_{\xi_I}(t),$$

where

$$\|\chi_{I,S}(t)\| \leq \bar{H}(z) [\|\eta_I(x_I(t), \bar{x}_I(t), u_I(t), t)\| + \|\Delta g_{I,S}(t)\|] + \|L_I\| \bar{\epsilon}_{\xi_I}(t).$$

In this case, it cannot be guaranteed that $\|\eta_I(x_I, \bar{x}_I, u_I, t)\| \leq \bar{\eta}_I(y_I(t), \bar{y}_{I,S}(t), u_I(t))$ (since Assumption 3 might not hold due to the sensor fault) or $\|\Delta g_{I,S}(t)\| \leq \bar{\Delta} g_I$ hold and as a result $\|\chi_{I,S}(t)\|$ may exceed $\bar{\chi}_I(t)$ given by (10) which is used by the local fault detection agent \mathcal{F}_I (note that in (10) the first term is actually $\bar{\eta}_I(y_I(t), \bar{y}_{I,S}(t), u_I(t))$ due to the use of the faulty measurements). Therefore, the residual may exceed its corresponding detection threshold, i.e. $|r_I^{(j)}(t)| > \bar{r}_I^{(j)}(t)$ for some $j = 1, \dots, p_I$ which means that the sensor fault that occurred in Σ_J can be detected by the local fault detection agent \mathcal{F}_I . \square

Remark 4: A qualitative explanation can be given for Lemma 2 as follows. In the case of a process fault that occurs in Σ_J , the fault affects its states which in turn affect other subsystems through the interconnection variables. So, the states of Σ_J are ‘‘contaminated’’ by the process fault and the measurements of (some of) these states also contain the process fault effects. Therefore, a subsystem Σ_I that is affected by Σ_J , is affected by the process fault that occurred in Σ_J through the interconnection variables $\bar{C}_I \bar{x}_I$ and the detection agent \mathcal{F}_I makes use of the measurements \bar{y}_I which are also ‘‘contaminated’’ by the same fault. Hence, the effect of the process fault that occurred in Σ_J , is ‘‘canceled out’’ in the detection agent \mathcal{F}_I and it is unable detect the fault. Hence, a process fault occurring in subsystem Σ_J is detectable only by its respective detection agent \mathcal{F}_J and not by any other detection agent \mathcal{F}_I . On the other hand, assume that a fault is detected by the detection agent \mathcal{F}_I and we know that it is a sensor fault. Then the faulty sensor might be due to the measurements y_I of Σ_I or due to the interconnection measurements \bar{y}_I of the other subsystems. This is because a sensor fault occurring in one subsystem affects the estimation model of its respective detection agent \mathcal{F}_I through y_I and also the estimation models of other agents \mathcal{F}_J through the communicated interconnection measurements \bar{y}_J , whereas the actual subsystems are influenced by the fault-free states (and not the faulty measurements).

Theorem 2 established a detectability condition for fault detection by the agent \mathcal{F}_I when a sensor fault occurs in Σ_I . The following Theorem, gives a detectability condition for fault detection by the agent \mathcal{F}_I when a sensor fault occurs in Σ_J which affects Σ_I .

Theorem 3 (Propagation Sensor Fault Detectability): *Consider the nonlinear interconnected system (1), (2) with the distributed fault detection scheme described in (4), (5), (6), (10) and (9). A sensor fault that occurs at $t = T_0^y$ in Σ_J which affects Σ_I , is detectable by the local fault detection agent \mathcal{F}_I if the following mismatch function $\Delta g_I''$ due to the sensor fault*

$$\Delta g_I''(t) \triangleq g_I(y_I^0(t), \bar{y}_I^0(t), u_I(t)) - g_I(y_I^0(t) + \xi_I(t), \bar{y}_I^0(t) + \bar{\xi}_I(t) + \bar{\beta}_I^y(t - T_0^y) \bar{\theta}_I(t), u_I(t))$$

satisfies the following inequality at some time t , for some $j = 1, 2, \dots, p_I$:

$$\left| \sum_{k=0}^{t-1} C_I^{(j)} A_{I,0}^{t-1-k} H(z) [\Delta g_I''(k)] \right| > \bar{r}_I^{(j)}(t) + \bar{\epsilon}_{\xi_I}(t) \\ + \sum_{k=0}^{t-1} \alpha_{I,j} \delta_{I,j}^{t-1-k} (\|H(z) [\eta_I(x_I(k), \bar{x}_I(k), u_I(k), k)]\| + \|L_I\| \bar{\epsilon}_{\xi_I}(k)). \quad (33)$$

Proof. See Appendix B. □

As it was shown in this Section, when a process fault occurs in subsystem Σ_I , then it can only be detected by its respective fault detection agent \mathcal{F}_I whereas when a sensor fault occurs in subsystem Σ_I , then it can be detected either by its respective detection agent \mathcal{F}_I or any other interconnected agent \mathcal{F}_J (monitoring Σ_J which is affected by Σ_I). This discriminating factor is exploited in the following Section for devising a high-level fault isolation scheme.

6. High-level fault isolation

In this section, we exploit the findings of the previous analysis for the derivation of a high-level fault isolation scheme (see Figure 1). The purpose of this high-level isolation scheme is to infer some conclusions regarding the type and/or location of the fault that has occurred in the whole interconnected system according to the decisions of the detection agents, although, it does not necessarily mean that exact fault identification/isolation can be achieved. As a result, the high-level isolation scheme can provide valuable information that can be used by a more advanced fault isolation scheme in order to greatly improve its performance by excluding potential fault scenarios.

In the analysis so far, we have considered the cases of process and sensor fault separately in order to gain some intuition of how these faults affect the local and neighboring detection agents. The main conclusion of this analysis was given in Lemma 2 which also constitutes the basis of the subsequent high-level isolation scheme. Although the proposed fault detection scheme may handle multiple faults, which can be both process and sensor faults, for the sake of the proposed high-level isolation scheme it is assumed that only one fault can occur among all subsystems, which may either be a process or a single sensor fault. Ideally, we would like to identify the type of fault that has occurred, that is, whether it is a process or sensor fault and furthermore in the case of a sensor fault to identify the faulty sensor.

In order to identify the sensors let us consider the following:

Definition 1: Let $\mathcal{S}\{y_I\}$ be the set of p_I sensors that measure $y_I \in \mathbb{R}^{p_I}$, $\mathcal{S}\{y_I\} \cap \mathcal{S}\{y_J\}$ be the set of common sensors among y_I and y_J and $\mathcal{S}\{y_I\} \cup \mathcal{S}\{y_J\}$ be the union of the set of sensors that measure y_I and y_J .

In addition let \mathcal{M} be the set of indices of the local fault detection agents that have detected a fault, i.e. $\mathcal{M} \triangleq \{I \in \{1, \dots, N\} : \mathcal{F}_I \text{ detects fault}\}$, let \mathcal{M}_i indicate the i -th index of the set \mathcal{M} and, let $m \in \{0, 1, \dots, N\}$ be the cardinality of the set \mathcal{M} , i.e. $m \triangleq \text{card}(\mathcal{M})$.

Then the following high-level isolation facts can be deduced:

- If $m = 0$ then all subsystems are considered as potentially non-faulty (the possibility of a fault not yet detected cannot be excluded).
- If $m = 1$ then the fault may be one of the following:
 - a) a process fault that has occurred in $\Sigma_{\mathcal{M}_1}$, OR
 - b) a single sensor fault in $\mathcal{S}\{y_{\mathcal{M}_1}\}$, OR
 - c) a single sensor fault in $\mathcal{S}\{\bar{y}_{\mathcal{M}_1}\}$.

- If $m \geq 2$ then the possibility of a process fault can be excluded and hence it is guaranteed that a sensor fault has occurred. The faulty sensor can be isolated within the set given by

$$\bigcap_{i=1}^m \left(\mathcal{S}\{y_{\mathcal{M}_i}\} \cup \mathcal{S}\{\bar{y}_{\mathcal{M}_i}\} \right). \quad (34)$$

If we consider that multiple faults can occur in the subsystems (i.e. multiple process and/or multiple sensor faults) then the isolation logic is modified as follows:

- If $m = 0$ then all subsystems are considered as potentially non-faulty.
- If $m \geq 1$ then the fault(s) may be:
 - a) a single/multiple process fault(s) that occurred in $\Sigma_{\mathcal{M}_i}, i \in \{1, \dots, m\}$, AND/OR
 - b) a single/multiple sensor fault(s) within the set given by

$$\bigcup_{i=1}^m \left(\mathcal{S}\{y_{\mathcal{M}_i}\} \cup \mathcal{S}\{\bar{y}_{\mathcal{M}_i}\} \right). \quad (35)$$

The proposed high-level fault isolation scheme in the case of joint process and sensor faults assuming that only one fault can occur, results in more constrained fault possibilities in comparison to the multiple fault case, in the sense that if two or more fault detection agents detect a fault then the occurrence of a process fault is excluded, and hence the fault is guaranteed to be a sensor fault contained in the set of sensors described by (34). If only one fault detection agent detects a fault then the fault type cannot be determined apart from the fact that the fault can either be a process or a sensor fault in the agents' respective subsystem, or a sensor fault in neighboring subsystems that affect the subsystem the fault has been detected in. In the case we consider multiple faults, the isolation results include more possibilities about fault occurrences but still provide useful information about the fault type and the set of that the faulty sensors are contained in.

These results prove to be more valuable in the case we consider that only one type of fault can occur (process or sensor), as it is the majority of the research conducted in the literature. By considering that only process faults can occur, that is by assuming that all sensors are healthy, then the detection of a fault by any fault detection agent guarantees that the fault has occurred in the respective subsystem that the particular agent is monitoring. Hence, partial fault isolation is achieved in the sense that the faulty subsystem is identified. Even if two or more agents detect a fault, then this means that in each of the respective subsystems a process fault has occurred. In the case we consider only sensor faults can occur (no process faults), then the faulty sensor(s) can be isolated within the set given by (34) in the case of a single sensor fault or by (35) in the case of multiple sensor faults. This additional information can be used by a more sophisticated fault isolation scheme in order to enhance its performance, by excluding potential fault scenarios. Moreover, the proposed distributed fault detection approach has significant benefits in comparison to a centralized approach since it encompasses important fault isolation characteristics.

In the proposed scheme, each detection agent provides a binary decision regarding the detection of a fault in the subsystem it monitors and, according to the decisions of all the agents, the high-level isolation scheme provides some information regarding the type and location of the fault that occurred. Moreover, some hypotheses can be stated regarding the status of presence of the fault, i.e. if the fault is permanent, intermittent or temporary. For instance, the successive threshold crossings, in the event of fault detection, can be interpreted as "the fault(s) is still present" but, the scheme cannot determine precisely the status of the fault. Of course, some hypotheses can be stated; according to the rate of successive threshold crossings and the frequency this behavior is observed. For instance, permanent faults, given that they are sufficiently large, will most probably cause the residuals to exceed their thresholds almost all the time after the initial detection or at least demonstrate rapid threshold crossings. On the other hand, temporary faults can be identified if after some time of the initial fault detection, the residuals fall and stay below their corresponding thresholds indicating

potentially healthy operation. Finally, intermittent faults behavior, will exhibit a mixture of the permanent and temporary faults behavior, with these two phases repeating successively. Specifically, during a period of time the residual will exceed its threshold like in the case of a permanent fault, and afterwards it will be followed by a period of time where the residuals stay below their thresholds. In general, the issue of distinguishing between permanent, temporary and intermittent faults, requires further investigation and is out of the scope of the present paper.

7. Simulation Results

In this section, we consider a numerical example based on a system of two inverted pendulums connected by a spring. The discrete time models of the two subsystems $I = 1, 2$ are obtained from a modified version of the continuous time version in Spooner and Passino (1999) by using a forward Euler discretization with a time step $T_s = 0.0001$ s and are given by

$$\begin{aligned}x_I^{(1)}(t+1) &= x_I^{(1)}(t) + T_s x_I^{(2)}(t) \\x_I^{(2)}(t+1) &= x_I^{(2)}(t) + T_s \left(f_I^{(2)}(t) + w_I^{(2)}(t) + \eta_I^{(2)}(t) \right) \\y_I(t) &= x_I^{(1)}(t) + \xi_I(t)\end{aligned}$$

where for the first subsystem the nominal and interconnection functions are given by:

$$\begin{aligned}f_1^{(2)}(t) &= \left(\frac{m_1 g r}{J_1} - \frac{k r^2}{4 J_1} \right) \sin(x_1^{(1)}(t)) + \frac{k r}{2 J_1} (l - b) + \frac{u_1}{J_1} \\w_1^{(2)}(t) &= \frac{k r^2}{4 J_1} \sin(x_2^{(1)}(t))\end{aligned}$$

and for the second subsystem the respective functions are

$$\begin{aligned}f_2^{(2)}(t) &= \left(\frac{m_2 g r}{J_2} - \frac{k r^2}{4 J_2} \right) \sin(x_2^{(1)}(t)) + \frac{k r}{2 J_2} (l - b) + \frac{u_2}{J_2} \\w_2^{(2)}(t) &= \frac{k r^2}{4 J_2} \sin(x_1^{(1)}(t)).\end{aligned}$$

The modification of this model with respect to Spooner and Passino (1999), is with regards to the availability of the state variables for measurement, the presence of modeling uncertainty and measurement noise. Specifically, in Spooner and Passino (1999) full state measurement is considered, whereas in this example it is considered that only $x_I^{(1)}(t)$, $I = 1, 2$ can be measured (with some uncertainty).

The parameters that are used in the simulation are: $m_1=2$ kg, $m_2=2.5$ kg, $J_1=0.5$ kg, $J_2=0.625$ kg, $k=30$ N/m, $l=0.5$ m, $b=0.4$ m and $g=9.81$ m/s². The modeling uncertainties of the subsystems are assumed to be $\eta_I^{(2)}(x_I^{(1)}, t) = 0.05 \sin(10t) + 0.05 \sin(x_I^{(1)}(t))$, $I = 1, 2$, in which the term $0.05 \sin(10t)$ corresponds to the uncertainty associated with time variations or general inaccuracies, whereas the term $0.05 \sin(x_I^{(1)}(t))$ corresponds to the uncertainty of the nominal function due to the error on some model parameters (i.e. the mass). Note that, the modeling uncertainty is a function of both the time t and the state $x_I^{(1)}$. The bound on the modeling uncertainty that is used is given by $\bar{\eta}_I^{(2)}(y_I^{(1)}, t) = 0.1 + 0.05 |\sin(y_I^{(1)}(t))|$, $I = 1, 2$, which satisfies Assumption 3. The inputs u_I are derived based on a simple decentralized proportional feedback controller that stabilizes each subsystem and are given by $u_I = 20e_I$, $I = 1, 2$ where $e_I = -y_I^{(1)}$ is the tracking error. In this example, we consider

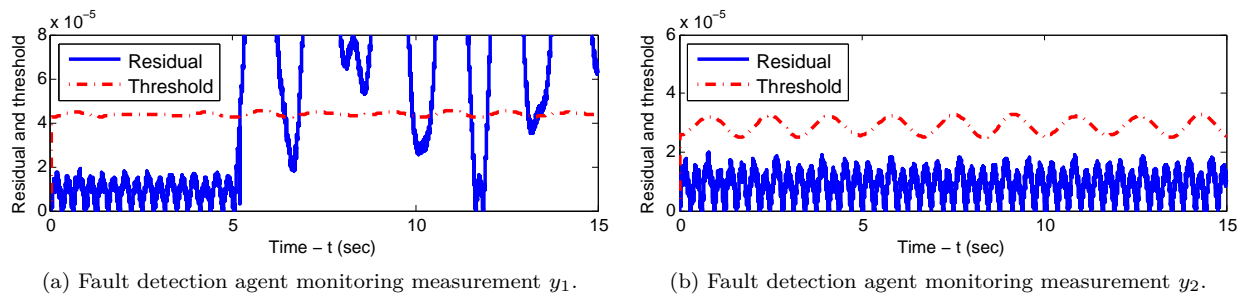


Figure 3.: Residual signal and fault detection threshold for measurements y_1 , y_2 in the case of process fault occurring in Σ_1 .

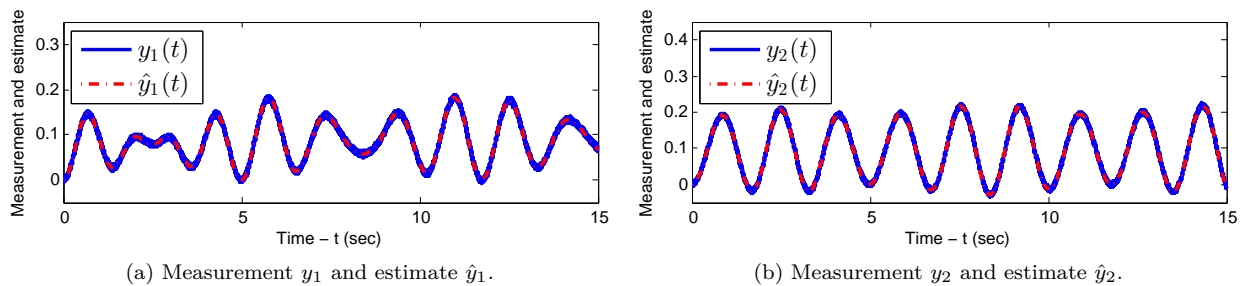


Figure 4.: Measurements y_1 , y_2 and their corresponding estimates \hat{y}_1 , \hat{y}_2 in the case of process fault occurring in Σ_1 .

two cases, one for process fault and one for the sensor fault. In the case of a process fault, we consider an abrupt multiplicative actuator fault in subsystem 1 where the input changes to $u_1 = (1 + \beta_1)\bar{u}_1$, where \bar{u}_1 is the nominal control input in the non-fault case and $\beta_1 \in [-1, 0]$ is the parameter characterizing the magnitude of the fault. The actuator fault in this case can be considered as a process fault affecting the dynamics of the system. The fault occurs at $T_0^x = 5$ sec with a magnitude $\beta_1 = -0.1$. In the case of a sensor fault, we consider that the sensor in the first subsystem measuring y_1 measures the signals' amplitude with 20% deviation and the sensor fault occurs at $T_0^y = 5$ sec.

The measurement noise ξ_I is implemented as a uniform random number in the range $[-0.01, 0.01]$. The proposed fault detection scheme is implemented using a FIR filter for $H(z)$. Specifically, the filter $H(z)$ is designed as a 10-th order FIR lowpass filter with normalized cutoff frequency 0.2 and utilizing a Hamming window (using the `fir1` command in Matlab). The transfer function of $H(z)$ is given by $H(z) = \sum_{k=0}^{10} d_k z^{-k}$ and explained in Section 3.3 the filter $\bar{H}(z)$ is given by $\bar{H}(z) = \sum_{k=0}^{10} |d_k| z^{-k}$. Using the aforementioned filter $H(z)$ the bounds on the filtered noise are found through the simulation as $\bar{\epsilon}_{\xi_I} = 4e-6$ and $\bar{\epsilon}_{\Delta g_I} = 1.2e-4$ for $I = 1, 2$.

According to the proposed fault detection scheme, two fault detection agents are designed, one for each subsystem. In each detection agent, the estimation model is given by (4), (5), the residual is generated according to (6) and finally, the detection threshold is generated according to (9) with (27) to fully exploit the filtering benefits. For the threshold implementation the constants $\alpha_{I,1} = 1$, $\delta_{I,1} = 0.7$, $I = 1, 2$, are also used so that (12) is satisfied. The simulation results for the fault detection agents that monitor the first subsystem (measured variable y_1) and the second subsystem (measured variable y_2) in the case of the process fault are shown in Figure 3 and in the case of the sensor fault in Figure 5.

At first let's consider the case of the process fault. In this case the measurements and their corresponding estimates of the two subsystems are shown in Figure 4 where it can be seen that the estimation model tracks the measurement although, it cannot be seen any significant discrepancy due to the process fault. The discrepancy of course is present, and it is the cause for the residual

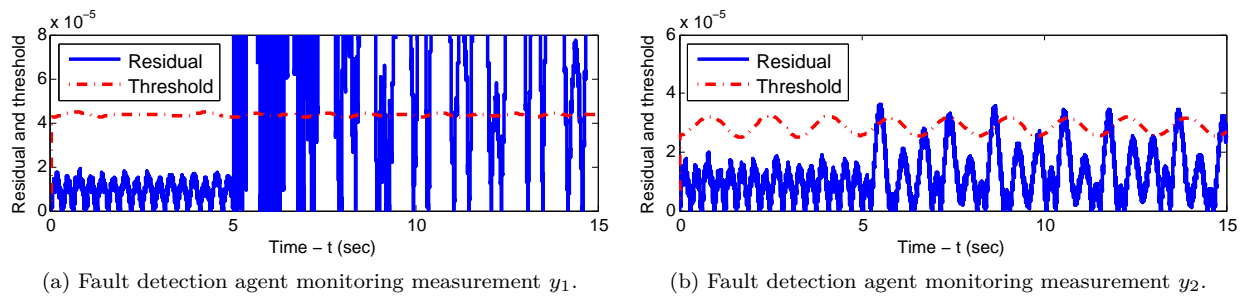


Figure 5.: Residual signal and fault detection threshold for measurements y_1 , y_2 in the case of sensor fault occurring in Σ_1 .

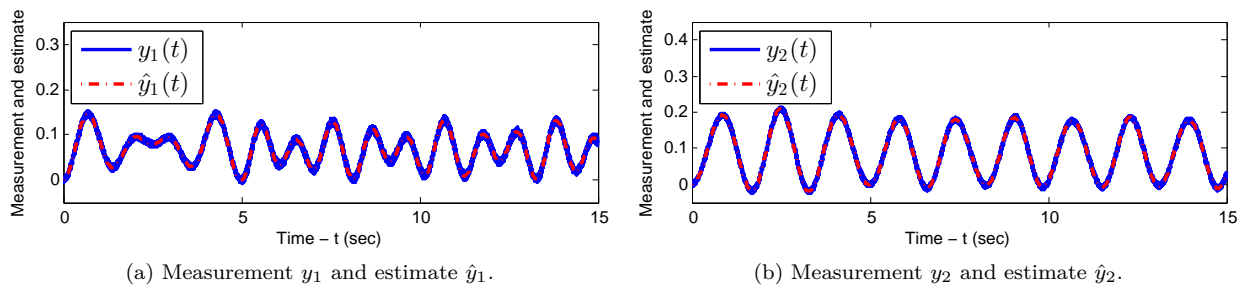


Figure 6.: Measurements y_1 , y_2 and their corresponding estimates \hat{y}_1 , \hat{y}_2 in the case of sensor fault occurring in Σ_1 .

signal given in Figure 3a to exceed its threshold. Specifically, the simulation results that correspond to the detection agent that monitors the measurements y_1 (1st subsystem) are shown in Figure 3a where it is clearly seen that the residual significantly exceeds its detection threshold after the fault occurs, hence the fault is detected at around $t = 5.09$ sec. It must be noted that there are no false alarms prior to the fault occurrence in either case. The corresponding results of the detection agent that monitors y_2 (2nd subsystem) are shown in Figure 3b where it is seen that the residual signal is always below its threshold signal and therefore no fault is detected. It must be noted that, the residual in the case of the second agent monitoring y_2 in Figure 3b, exhibits the same behavior before and after the occurrence of the fault, indicating that the process fault effects from the first subsystem do not seem to impact the residual of the second agent.

Now let us consider the case of the sensor fault. In this case, the measurements and their corresponding estimates of the two subsystems are shown in Figure 6. In the case of the sensor fault, the results in the case of the fault detection agents that monitor the measurements y_1 and y_2 are shown in Figure 5a and 5b respectively. In this case, the sensor fault is detected very fast by the detection module that monitors y_1 at around $t = 5.01$ sec. Most importantly though, the sensor fault is also detected by the detection module that monitors y_2 at around $t = 5.39$ sec. As in the previous case, no false alarms occur, since the residuals are always bounded by their thresholds prior to the sensor fault occurrence. In this case, it must be noted that, the residual behavior in the case of the second agent monitoring y_2 in Figure 5b, changes after the occurrence of the fault, indicating that the sensor fault effects from the first subsystem affect significantly the residual of the second agent.

The simulation results are in line with the findings of the analysis conducted in Section 4 and more specifically with Lemma 2 according to which the process fault occurring in a subsystem can only be detected by its respective fault detection agent, whereas a sensor fault occurring in a subsystem can be detected by its respective fault detection and also by neighboring interconnected detection agents. According to the high-level fault isolation scheme described in Section 6, in the case of the process fault, no conclusive decision can be reached since only one detection agent detects a fault

and therefore the fault can be either a process fault in Σ_1 (which is actually the case), or a fault in the sensors measuring y_1 or y_2 . In the case of the sensor fault, because both fault detection agents detect a fault, the case of a process fault is excluded and hence it is guaranteed that a sensor fault has occurred and the faulty sensor is either y_1 (which is actually the case) or y_2 . Therefore, further actions can be taken to identify precisely the type of the fault that has occurred.

8. Conclusion

In this paper, a distributed fault diagnosis approach for the detection of process and sensor faults in a class of interconnected input-output discrete-time, nonlinear systems with modeling uncertainties and measurement noise is presented. By utilizing a filtering approach which is incorporated in the fault detection framework to mitigate the measurement noise effects, robust adaptive thresholds are designed that guarantee no false alarms. Furthermore, the propagation of a fault that occurs in one subsystem and affects neighboring subsystems is investigated, leading to some key properties of fault propagation among subsystems. More specifically, the fault detection scheme is designed in such a way that a process fault occurring in a subsystem can only be detected by its corresponding detection agent, whereas a sensor fault occurring in a subsystem can also be detected by the detection agents of the neighboring subsystems it affects. This discriminating element is exploited to extrapolate further information regarding the type of fault that has occurred and, constitutes the basis of the derived high-level isolation scheme. Furthermore, detectability conditions have been derived that characterize quantitatively the class of process and sensor faults that can be detected by the proposed scheme. Future research efforts will be devoted in the development of a comprehensive fault isolation methodology and, the integration of filtering with learning techniques in order to derive tighter detection thresholds by dampening the measurement noise and by learning the modeling uncertainty.

Appendix A. Proof of Theorem 2

In the presence of a sensor fault that occurs in the I -th subsystem at $t = T_0^y$, (20) becomes

$$\tilde{x}_{I,f}(t) = A_{I,0}^t \tilde{x}_{I,f}(0) + \sum_{k=0}^{t-1} A_{I,0}^{t-1-k} \chi_I'(k), \quad (\text{A1})$$

where

$$\chi_I'(t) \triangleq H(z)[\eta_I(x_I(t), \bar{x}_I(t), u_I(t), t) + \Delta g_I'(t)] - L_I \epsilon_{\xi_I}(t) - L_I \theta_{I,f}(t).$$

Let's define $v_{I,f}(t) \triangleq H(z)[\Delta g_I'(t)] - L_I \theta_{I,f}(t)$. After the occurrence of a sensor fault, the residual (6) becomes $r_I(t) = C_I \tilde{x}_{I,f}(t) + \epsilon_{\xi_I}(t) + \theta_{I,f}(t)$ and by using (A1) with $\tilde{x}_{I,f}(0) = 0$, the residual is written as:

$$\begin{aligned} r_I(t) &= \sum_{k=0}^{t-1} C_I A_{I,0}^{t-1-k} \chi_I'(k) + \epsilon_{\xi_I}(t) + \theta_{I,f}(t) \\ &= \sum_{k=0}^{t-1} C_I A_{I,0}^{t-1-k} (H(z)[\eta_I(x_I(k), \bar{x}_I(k), u_I(k), k)] - L_I \epsilon_{\xi_I}(k)) \\ &\quad + \sum_{k=0}^{t-1} C_I A_{I,0}^{t-1-k} v_{I,f}(k) + \epsilon_{\xi_I}(t) + \theta_{I,f}(t) \end{aligned}$$

By using the triangle inequality, the j -th element of $r_I(t)$ for $t > T_0^y$ satisfies:

$$\begin{aligned}
 |r_I^{(j)}(t)| &\geq \left| \theta_{I,f}^{(j)}(t) + \sum_{k=0}^{t-1} C_I^{(j)} A_{I,0}^{t-1-k} v_{I,f}(k) \right| - |\epsilon_{\xi_I}^{(j)}(t)| \\
 &\quad - \left| \sum_{k=0}^{t-1} C_I^{(j)} A_{I,0}^{t-1-k} (H(z)[\eta_I(x_I(k), \bar{x}_I(k), u_I(k), k)] - L_I \epsilon_{\xi_I}(k)) \right| \\
 &\geq \left| \theta_{I,f}^{(j)}(t) + \sum_{k=0}^{t-1} C_I^{(j)} A_{I,0}^{t-1-k} v_{I,f}(k) \right| - \bar{\epsilon}_{\xi_I}(t) \\
 &\quad - \sum_{k=0}^{t-1} \alpha_{I,j} \delta_{I,j}^{t-1-k} (\|H(z)[\eta_I(x_I(k), \bar{x}_I(k), u_I(k), k)]\| + \|L_I\| \bar{\epsilon}_{\xi_I}(k)),
 \end{aligned}$$

For fault detection, the inequality $|r_I^{(j)}(t)| > \bar{r}_I^{(j)}(t)$ must hold for some $j = 1, 2, \dots, p_I$, so the final fault detectability condition given in (30) is obtained.

Appendix B. Proof of Theorem 3

In the presence of a sensor fault that occurs at $t = T_0^y$ in Σ_J which affects Σ_I , the state estimation error (20) becomes

$$\tilde{x}_{I,f}(t) = A_{I,0}^t \tilde{x}_{I,f}(0) + \sum_{k=0}^{t-1} A_{I,0}^{t-1-k} \chi_I''(k), \tag{B1}$$

where

$$\chi_I''(t) \triangleq H(z)[\eta_I(x_I(t), \bar{x}_I(t), u_I(t), t) + \Delta g_I''(t)] - L_I \epsilon_{\xi_I}(t).$$

After the occurrence of a sensor fault in Σ_J , the residual (6) becomes $r_I(t) = C_I \tilde{x}_{I,f}(t) + \epsilon_{\xi_I}(t)$ and by using (B1) with $\tilde{x}_{I,f}(0) = 0$ due to the filters' initial condition, the residual is written as:

$$\begin{aligned}
 r_I(t) &= \sum_{k=0}^{t-1} C_I A_{I,0}^{t-1-k} \chi_I''(k) + \epsilon_{\xi_I}(t) \\
 &= \sum_{k=0}^{t-1} C_I A_{I,0}^{t-1-k} H(z)[\Delta g_I''(k)] + \epsilon_{\xi_I}(t) \\
 &\quad + \sum_{k=0}^{t-1} C_I A_{I,0}^{t-1-k} (H(z)[\eta_I(x_I(k), \bar{x}_I(k), u_I(k), k)] - L_I \epsilon_{\xi_I}(k))
 \end{aligned}$$

By using the triangle inequality, the j -th element of $r_I(t)$ for $t > T_0^y$ satisfies:

$$\begin{aligned}
|r_I^{(j)}(t)| &\geq \left| \sum_{k=0}^{t-1} C_I^{(j)} A_{I,0}^{t-1-k} H(z) [\Delta g_I''(k)] \right| - |\epsilon_{\xi_I}^{(j)}(t)| \\
&\quad - \left| \sum_{k=0}^{t-1} C_I^{(j)} A_{I,0}^{t-1-k} (H(z) [\eta_I(x_I(k), \bar{x}_I(k), u_I(k), k)] - L_I \epsilon_{\xi_I}(k)) \right| \\
&\geq \left| \sum_{k=0}^{t-1} C_I^{(j)} A_{I,0}^{t-1-k} H(z) [\Delta g_I''(k)] \right| - \bar{\epsilon}_{\xi_I}(t) \\
&\quad - \sum_{k=0}^{t-1} \alpha_{I,j} \delta_{I,j}^{t-1-k} (\|H(z) [\eta_I(x_I(k), \bar{x}_I(k), u_I(k), k)]\| + \|L_I\| \bar{\epsilon}_{\xi_I}(k)).
\end{aligned}$$

For fault detection, the inequality $|r_I^{(j)}(t)| > \bar{r}_I^{(j)}(t)$ must hold for some $j = 1, 2, \dots, p_I$, so the final fault detectability condition given in (33) is obtained.

Acknowledgements

This work was supported by funding from the European Research Council under the ERC Advanced Grant (FAULT-ADAPTIVE).

References

- Basseville, M., & Nikiforov, I. (1993). *Detection of abrupt changes: theory and application*. Prentice-Hall.
- Blanke, M., Kinnaert, M., Lunze, J., & Staroswiecki, M. (2010). *Diagnosis and Fault-Tolerant Control* (2nd ed.). Springer Verlag.
- Boem, F., Ferrari, R. M., & Parisini, T. (2011). Distributed Fault Detection and Isolation of Continuous-Time Nonlinear Systems. *European Journal of Control*, 5-6, 603–620.
- Chen, J., & Patton, R. J. (1999). *Robust Model-Based Fault Diagnosis for Dynamic Systems*. Kluwer Academic Publishers Norwell, MA, USA.
- De Persis, C., & Isidori, A. (2002). On the design of fault detection filters with game-theoretic-optimal sensitivity. *International Journal of Robust and Nonlinear Control*, 12(8), 729–747.
- Dunia, R., & Joe Qin, S. (1998). Joint diagnosis of process and sensor faults using principal component analysis. *Control Engineering Practice*, 6(4), 457–469.
- Ferdowsi, H., Raja, D., & Jagannathan, S. (2012). A decentralized fault prognosis scheme for nonlinear interconnected discrete-time systems. In *American Control Conference* (pp. 5900–5905).
- Ferrari, R. M., Parisini, T., & Polycarpou, M. M. (2008). A robust fault detection and isolation scheme for a class of uncertain input-output discrete-time nonlinear systems. In *American Control Conference* (pp. 2804–2809).
- Ferrari, R. M., Parisini, T., & Polycarpou, M. M. (2012). Distributed fault detection and isolation of large-scale nonlinear systems: an adaptive approximation approach. *IEEE Transactions on Automatic Control*, 57(2), 275–290.
- Frank, P. (1990). Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy: A survey and some new results. *Automatica*, 26(3), 459–474.
- Hammouri, H., Kinnaert, M., & El Yaagoubi, E. (1999). Observer-based approach to fault detection and isolation for nonlinear systems. *IEEE Transactions on Automatic Control*, 44(10), 1879–1884.
- Keliris, C., Polycarpou, M. M., & Parisini, T. (2013a). A Distributed Fault Detection Filtering Approach for a Class of Interconnected Continuous-Time Nonlinear Systems. *IEEE Transactions on Automatic Control*, 58(8), 2032–2047.

- Keliris, C., Polycarpou, M. M., & Parisini, T. (2013b). A Distributed Fault Detection Filtering Approach for a Class of Interconnected Input-Output Nonlinear Systems. In *European Control Conference* (pp. 422–427). Zurich.
- Kinnaert, M., & Peng, Y. (1995). Residual generator for sensor and actuator fault detection and isolation: a frequency domain approach. *International Journal of Control*, 61(6), 1423–1435.
- Klinkhieo, S., & Patton, R. J. (2009). A Two-Level Approach to Fault-Tolerant Control of Distributed Systems Based on the Sliding Mode. In *7th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes, Barcelona, Spain* (pp. 1043–1048).
- Léchevin, N., & Rabbath, C. (2009). Decentralized Detection of a Class of Non-Abrupt Faults With Application to Formations of Unmanned Airships. *IEEE Transactions on Control Systems Technology*, 17(2), 484–493.
- Liu, Y., & Bauer, P. (2008). Sufficient conditions for non-negative impulse response of arbitrary-order systems. In *IEEE Asia Pacific Conference on Circuits and Systems* (pp. 1410–1413).
- Massoumnia, M.-A., & Vander Velder, W. E. (1988). Generating parity relations for detecting and identifying control system component failures. *Journal of Guidance, Control, and Dynamics*, 11(1), 65.
- Patton, R., & Frank, P. M. (1989). *Fault diagnosis in dynamic systems: theory and applications*. Prentice Hall.
- Patton, R. J., Kambhampati, C., Casavola, A., Zhang, P., Ding, S. X., & Sauter, D. (2007). A generic strategy for fault-tolerance in control systems distributed over a network. *European Journal of Control*, 13(2-3), 280–296.
- Rajamani, R., & Ganguli, A. (2004). Sensor fault diagnostics for a class of non-linear systems using linear matrix inequalities. *International Journal of Control*, 77(10), 920–930.
- Reppa, V., Polycarpou, M., & Panayiotou, C. (2014). Adaptive Approximation for Multiple Sensor Fault Detection and Isolation of Nonlinear Uncertain Systems. *IEEE Transactions on Neural Networks and Learning Systems*, 25(1), 137–153.
- Salahshoor, K., Mosallaei, M., & Bayat, M. (2008). Centralized and decentralized process and sensor fault monitoring using data fusion based on adaptive extended Kalman filter algorithm. *Measurement*, 41(10), 1059–1076.
- Spooner, J., & Passino, K. (1999). Decentralized adaptive control of nonlinear systems using radial basis neural networks. *IEEE Transactions on Automatic Control*, 44(11), 2050–2057.
- Stankovic, S., Ilic, N., Djurovic, Z., Stankovic, M., & Johansson, K. (2010). Consensus based overlapping decentralized fault detection and isolation. In *Conference on Control and Fault-Tolerant Systems (SysTol'10)*, (pp. 570–575).
- Talebi, H. a., Khorasani, K., & Tafazoli, S. (2009). A recurrent neural-network-based sensor and actuator fault detection and isolation for nonlinear systems with application to the satellite's attitude control subsystem. *IEEE Transactions on Neural Networks*, 20(1), 45–60.
- Thumati, B., & Halligan, G. (2013). A Novel Fault Diagnostics and Prediction Scheme Using a Nonlinear Observer With Artificial Immune System as an Online Approximator. *IEEE Transactions on Control Systems Technology*, 21(3), 569–578.
- Venkatasubramanian, V., Rengaswamy, R., Yin, K., & Kavuri, S. (2003, March). A review of process fault detection and diagnosis Part I: Quantitative model-based methods. *Computers & Chemical Engineering*, 27(3), 293–311.
- Viswanadham, N., & Srichander, R. (1987). Fault detection using unknown-input observers. *Control Theory and Advanced Technology*, 3(2), 91–101.
- Wei, L., Gui, W., Xie, Y., & Ding, S. X. (2009). Decentralized Fault Detection System Design for Large-Scale Interconnected Systems. In *7th IFAC symposium on Fault Detection, Supervision and Safety of Technical Processes, Barcelona, Spain* (pp. 816–821).
- Yan, B., Tian, Z., & Shi, S. (2008). A novel distributed approach to robust fault detection and identification. *International Journal of Electrical Power & Energy Systems*, 30(5), 343–360.
- Zhang, Q., Basseville, M., & Benveniste, A. (1998). Fault detection and isolation in nonlinear dynamic systems: A combined input-output and local approach. *Automatica*, 34(11), 1359–1373.
- Zhang, Q., & Zhang, X. (2012). A distributed detection scheme for process faults and sensor faults in a class of interconnected nonlinear uncertain systems. In *IEEE 51st Annual Conference on Decision and Control* (pp. 586–591).
- Zhang, X., Polycarpou, M. M., & Parisini, T. (2002). A robust detection and isolation scheme for abrupt and incipient faults in nonlinear systems. *IEEE Transactions on Automatic Control*, 47(4), 576–593.

- Zhang, X., Polycarpou, M. M., & Parisini, T. (2008). Design and analysis of a fault isolation scheme for a class of uncertain nonlinear systems. *Annual Reviews in Control*, 32(1), 107–121.
- Zhang, X., Polycarpou, M. M., & Parisini, T. (2009). Decentralized fault detection for a class of large-scale nonlinear uncertain systems. In *48th IEEE Conference on Decision and Control and 28th Chinese Control Conference* (pp. 6988–6993).