Chapter 1
Fault Diagnosis for Uncertain Networked Systems [*]

Francesca Boem, Christodoulos Keliris, Thomas Parisini and Marios M. Polycarpou

Abstract Fault diagnosis has been at the forefront of technological developments for several decades. Recent advances in many engineering fields have led to the networked interconnection of various systems. The increased complexity of modern systems leads to a larger number of sources of uncertainty which must be taken into consideration and addressed properly in the design of monitoring and fault diagnosis architectures. This chapter reviews a model-based distributed fault diagnosis approach for uncertain nonlinear large-scale networked systems to specifically address: a) the presence of measurement noise by devising a filtering scheme for dampening the effect of noise; b) the modeling of uncertainty by developing an adaptive learning scheme; c) the uncertainty issues emerging when considering networked systems, such as the presence of delays and packet dropouts in the communication networks. The proposed architecture considers in an integrated way the various components of complex distributed systems, such as: the physical environment, the sensor level, the fault diagnosers and the communication networks. Finally, some actions taken after the detection of a fault, such as the identification of the fault location and its magnitude or the learning of the fault function, are illustrated.

Francesca Boem
University College London, Torrington Place, WC1E 7JE London, UKe-mail: f.boem@ucl.ac.uk

Christodoulos Keliris
KIOS Research and Innovation Center of Excellence, University of Cyprus, 1 Panepistimiou Avenue, Aglantzia, CY-2109 Nicosia, Cyprus, e-mail: keliris.chris@gmail.com

Thomas Parisini
Imperial College London, KIOS Research and Innovation Center of Excellence, and University of Trieste e-mail: t.parisini@gmail.com

Marios M. Polycarpou
KIOS Research and Innovation Center of Excellence, University of Cyprus, 1 Panepistimiou Avenue, Aglantzia, CY-2109 Nicosia, Cyprus, e-mail: mpolycar@ucy.ac.cy

[*] This work has been supported by the European Union's Horizon 2020 research and innovation programme under grant agreement No 739551 (KIOS CoE).

1.1 Introduction: from centralised to distributed fault diagnosis

In systems and control engineering, the adoption of models describing the behaviour of systems is ubiquitous and of fundamental importance. However, such models are usually affected by some uncertainty and, the sources of uncertainty may vary quite a lot. For instance, the derivation of an accurate mathematical model may be very difficult to obtain or even entail increased financial costs and so, less accurate models are used. Other sources of uncertainty include the measurement noise, the system disturbances and the changing system parameters due to the components degradation over time. The presence of uncertainty is especially important when considering complex large-scale systems, such as Systems-of-Systems (SoS) [79] or Cyber Physical Systems (CPS) [4], where it is difficult to understand and model the relationships that exist among the (possibly large) number of interconnected subsystems. Therefore, uncertainty represents an important challenge for many control applications, thus motivating the research and the development of robust methods able to manage its presence and effect on the control task performance [25, 109, 97, 67]. In some situations, the mismatch between the considered model and the actual system behavior becomes major, due to the presence of undesired or unexpected behaviours, possibly leading to negative consequences, such as instabilities, failures in the system or deterioration of performance. Therefore, it is important to take into consideration modeling uncertainty at the design stage, so that if any unexpected behaviour is observed during the system operation, it will be feasible to identify the presence of a fault, avoiding, at the same time, the occurrence of false-alarms.

Reliability is a key requirement for modern systems. It can be defined as the ability of a system to perform its intended function over a given period of time [7]. The inability to perform the intended function is called a failure, and it can be due to the effects of a fault. A fault is a change in the behavior of a system, or part of it, from the behavior that was set at design time.

As practical systems become more complex and more interconnected, the need for enhanced robustness, fault tolerance and sustainability becomes of essential importance. Potential faults could lead to major catastrophes and consequently could trigger a chain of failing dependent systems such as electric power systems, communication and water networks, along with production plants causing tremendous economic and social damage. Therefore, safe and reliable operation of such systems through the early detection of any “small” fault before they become serious failures is a crucial component of the overall system performance and sustainability.

For these reasons, fault diagnosis is a research field that has been in the forefront of the technological evolution for a few decades and has attracted the attention from the research and industrial communities, as testified by many important survey papers [33, 37, 100, 101, 99, 43] and books [9, 18, 65, 44].

Generally, fault diagnosis is comprised of several steps: *detection* of a fault, *isolation* and *identification* of the fault and fault *accommodation* or reconfiguration of the system.

Fault detection consists of understanding whether a fault has occurred or not, while the isolation task refers to pinpointing the type of fault and its location. Fault identification is an extra step that is carried on after isolation in order to quantify the extent to which a fault is present. Fault accommodation addresses the problem of how the system actively responds to the fault: for example, after a successful fault diagnosis, the controller parameters may be adjusted to accommodate changed plant dynamics in order to prevent failure at the system level.

A control system is comprised of mainly three parts: the actuators, the plant components and the sensors, therefore a fault may appear in any of these (see Figure 1.1). Specifically, process faults (on the plant components) alter the dynamics of the system, sensor faults alter the measurement readings and actuator faults modify the controllers' influence on the system.

Apart from the fault source, we can further distinguish between abrupt or incipient faults. *Abrupt faults* are sudden, step-like changes that appear almost instantaneously and can lead to immediate component or even general system failure. On the other hand, *incipient faults* are slowly developing faults that occur due to parameter changes of the components because of their continuous operation and diminishing lifetime. These changes develop slowly and are initially small, thus harder to detect and may be better prevented through system maintenance.

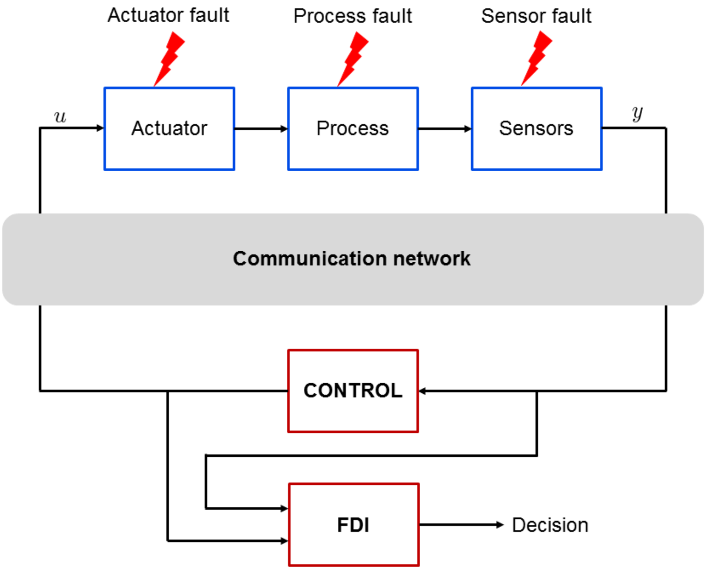

Fig. 1.1: Fault types and FDI.

There are mainly two methods to address the possible presence of a fault. The first one is *physical redundancy* (or hardware redundancy), that is the fact that critical components of the system are replicated in a greater number than what is strictly necessary. This is effective but implies a highly expensive solution and can be justified only for critical, potentially life-threatening systems (i.e. aviation applications). The second method is the *analytical redundancy* approach which is based on a mathematical model of the system under healthy system behavior. In this approach, the actual physical signals that are measured, are compared to the corresponding signals given by the mathematical model of the process under healthy state; their difference constitutes the *residuals* (*residual generation stage*). Under the ideal conditions of no faults, no modeling uncertainties and no measurement noise nor disturbances, the residuals are zero. In real applications, after the residual generation stage, the information given by residuals is processed to take a decision regarding the healthy status of the system and determine the potential occurrence of faults (*decision making stage*). If the fault decision is positive, then further analysis is conducted to identify the faults' type and location, and possibly its size. Although this approach is more affordable, it is computationally intensive and may be sensitive to false alarms due to inaccuracies in the mathematical modeling of the system which may be mistakenly passed as faults. This model-based approach was born during the 1970s thanks to the seminal works of Beard, Jones and Clark [5, 47, 22] among others (see the survey papers [37, 45, 33, 100]).

An alternative approach to model-based methods is represented by the signal-based techniques, in which known features of signals, such as spectral components or statistical features, are compared to nominal ones [44, 37]. These methods though, require some knowledge of previous behavior of the system during healthy operation and that is the reason they are classified into the wider class of process history fault diagnosis approaches (i.e. see [99] and the references therein).

Under the analytical redundancy framework, there are various methods to generate the residual vector, which can be divided into two main approaches: the state estimation techniques (such as parity space approach, observer schemes and detection filters) and the parameter identification techniques. Moreover, in order to ease the fault isolation task, residuals can be designed so as to contain specific isolation properties. The main residual enhancement techniques are represented by structured and directional residuals [100, 38]. In the *structured residuals* scheme, each fault affects a specific subset of the residuals and any residual responds only to a specific subset of faults. Therefore, due to the dependence of the residuals on the faults, certain patterns appear on the residual vector that can be used for fault isolation. In the *directional residuals* scheme, each fault amounts to a specific direction in the residual space, and thus fault isolation is concluded by selecting the direction that the generated residual vector lies closest to. More information regarding these techniques can be found in the books by Gertler [39] and Isermann [44]. In the literature, many methods have been proposed for the generation of residuals, which can mainly be classified according the following approaches:

- *Parity space approach*. This method consists of checking the consistency of the mathematical equations by using the actual measurements: a fault is declared whenever predetermined error thresholds are exceeded. Further information can be found in [38] and the references therein.

- *Observer schemes*. In this category lie many approaches, starting from the *Fault detection filter* (FDF), firstly proposed by Beard and Jones in the early 70's, to the *Diagnostic Observer* approach, which has been widely adopted in the literature. According to this approach, observers are used to reconstruct the output $\hat{y}$ of the system from measurements y and the residual is represented by the output estimation error $e = y - \hat{y}$. In the case of stochastic systems, the observers may be substituted by Kalman filters and the residual is the innovation which under the fault free case should be white noise with zero mean and known covariance. The isolation of faults can be enhanced with the use of a bank of residual generators under the *Dedicated Observer Scheme (DOS)* proposed by Clark [22] or the *Generalized Observer Scheme (GOS)* [33, 34]. In both schemes as many residuals as the number of possible faults are generated. The difference is that in the DOS scheme, each residual is sensitive to only a single fault, while in the GOS, each residual is sensitive to every but one fault. The DOS scheme is appealing as it can also isolate concurrent faults, but it cannot always be designed. Instead the GOS can be always applied, but can only isolate non-concurrent faults. It is important to note that, as pointed out in [34], the observers used in fault diagnosis are primarily output observers which simply reconstruct the measurable part of the state variables, rather than state observers which are required for control purposes. The use of state observers for nonlinear systems has not been used extensively for the FDI problem, even though analytical results regarding the stability of the nonlinear observers and design procedures have been established. The main issue with the observer approach is that the design of observers for nonlinear systems with asymptotically stable error dynamics is not an easy task even when the nonlinearities are fully known. As a result, the research in fault diagnosis for nonlinear systems utilizing state observers is more limited [36, 1, 41, 51].
- *Parameter estimation*. This method is particularly suited to the detection of incipient faults and it is extensively studied in the survey papers by Isermann [45] and Frank [33] and the books by Patton *et al.* [65] and Isermann [44]. Using system identification methods (utilizing the input and output signals), the parameters of a mathematical model of the system can be obtained (recursively and on-line) across different time intervals and compared to their respective values based on a nominal model. Any significant difference could indicate the occurrence of a fault and, a relation between parameter changes and faults can be formed with the use of pattern recognition methods.

An important aspect to be considered when monitoring controlled systems relates to the possibly *conflicting* dynamic behaviors of the FDI scheme and the reconfigurable controller, namely the feedback controller may hide the presence of faults by compensating their effects (see as example the simulation analysis in [78]) thus making the FDI task much more difficult or even impossible [3, 21, 100, 35]. This is particularly eminent in passive FDI methods, in which the status of health of the system is analyzed by comparing input-output data for the closed-loop system with a process model or historical data. A possible solution has been proposed for this problem when considering application use-cases allowing to affect the closed loop dynamics by acting at run-time on the control inputs. This paves the way to the so-called *active FDI methodologies*. Active FDI approaches consist of suitably modifying the control input to improve fault detectability and isolability capabilities [2, 71, 20, 82, 92, 87, 42, 6, 73]. The typical main limitation of active FDI techniques concerns high computational cost and complexity. This drawback restricts quite a bit the applicability of this approach to low-dimensional systems [86, 73, 85, 104, 30, 105], even though some approaches have been suggested in the literature to alleviate the computational complexity (see as examples [6, 62]).

An obvious problem in the practical implementation of model-based FDI schemes consists of deriving accurate mathematical models of engineering systems. This is a challenging task and thus, due to the presence of uncertainties and modelling errors, the resulting residual vectors are never identically zero. In addition, generally in the literature, the presence of measurement noise and modeling uncertainty is often overseen. In most real world applications such uncertainties may influence significantly the performance of fault detection schemes by causing, for example, false alarms. Therefore, bounds on the residuals must be defined, but still the proper choice remains a major problem. If bounds are chosen too narrow, this may lead to false alarms, whilst if they are chosen too wide faults may pass undetected. Therefore, dealing with the uncertainty in Fault Detection and Isolation architectures is of fundamental importance. As a result, there is a growing demand for robust residual generation to reduce the sensitivity of the residual against the effect of modeling errors, noise and disturbances. This issue can be tackled either by the use of enhanced techniques for robust residual generation or by choosing appropriately the level of the error threshold which can also change adaptively as discussed in the book by Patton *et al.* [65]. A line of research tried to overcome the problem of accurate mathematical modeling by using qualitative models, where only qualitative information, such as sign or trend of measured variables, are used [101] as well as classification techniques and inference methods. A more successful approach, anyway, is based on the use of adaptive on-line approximators, such as neural networks as example, to learn on-line the unknown or uncertain parts of the system dynamical model or the fault model [28, 69, 98, 107, 15, 31, 16, 53, 50].

1.1.1 Distributed and networked large-scale systems

In the literature, FDI methods have been historically designed for *centralized* frameworks, where information about the state of the system is gathered and processed centrally. From a practical perspective, gathering the distributed information into a central processing unit to implement a centralized approach for the fault diagnosis task is counter-productive due to communication overload and the requirement for higher computational power. Moreover the processing of the information at a centralized station imposes several risks since the station constitutes a *single point of failure*, thus making the architecture possibly fragile to faults. Recent advances in communications and distributed sensing have allowed the transition from centralized fault diagnosis approaches [33, 100, 18, 65, 9] towards the development of hierarchical, decentralized and distributed schemes [66, 54, 102, 56, 108, 89, 15, 84, 29, 31, 48, 49, 96, 90, 76, 52, 75, 55, 8, 40, 23, 78, 14, 13].

In many cases, a distributed FDI framework is not an option but a necessity, since many factors contribute to this formulation such as the large scale nature of the system to be monitored, its spatial distribution, the inability to access certain parts of the system from a remote monitoring component. Specifically, recent research efforts are focused on decentralized, distributed, networked systems, Cyber Physical Systems (CPS) [4] and Systems of Systems (SoS) [80]. Examples of these systems include power networks, water distribution networks, transportation systems, smart buildings and complex industrial plants. The term CPS refers to systems with integrated computational and physical capabilities that can interact with humans through many new modalities [4], expanding the capabilities of the physical world through computation, communication, and control. On the other hand, a SoS can be considered as a composition, made of components that are themselves systems, which is characterized by two properties that the whole must possess for it [61]: operational and managerial independence of components. This means that the component systems fulfill their own purposes and continue to operate to fulfill those purposes even if disassembled from the overall system; besides, the component systems are managed (at least in part) for their own purposes rather than the purposes of the whole.

In this chapter, we will use the term *networked* with two meanings: the considered system can be represented as a network of physically interconnected subsystems, and the monitoring agents operate and collaborate using input-output information obtained through a communication network.

When monitoring this kind of systems, distributed or decentralized algorithms are usually necessary due to computational, communication, scalability and reliability limits. The main benefits of using a distributed fault diagnosis scheme can be summarized as follows: a) enhanced robustness of the monitoring architecture, since centralized approaches are subject to single-point-of-failure, b) reduced computation costs, c) scalability benefits; the distributed scheme allows for more flexibility in adding subsystems with respective fault detection modules requiring fewer and possibly local modifications in the already existing architecture. Moreover, an emerging requirement is the design of monitoring architectures that are robust to changes that may occur in the dynamic topology of the large scale systems, allowing the addition/disconnection of subsystem to/from the network of interconnected subsystems only requiring local operations (see for example [78, 11, 13]).

Concerning Cyber-Physical Systems, in the literature many contributions deal with the description of the technical challenges and design and modeling issues that need to be addressed in order to interface with these modern systems, the technological impact deriving by CPS and the requirements emerging by them ([4] and [93, 58, 57, 83, 103, 74, 46, 59, 106]). With regards to reliability, safety and security of CPS, some methods have been proposed ([77], including some recent works dealing with the topic of the detection of cyber-physical attacks and attacks against process control systems [17, 95, 84, 26, 63, 64, 81, 88, 19]. An interesting approach for distributed fault diagnosis is based on exploiting sensor networks [32, 110].

Another important direction of research related to the control and monitoring of large-scale distributed networked systems is the design of distributed fault-tolerant control (FTC) architectures based on passive [8, 10, 91, 78] or active FDI methods [72].

1.1.2 Outline of the chapter

Motivated by the issues raised above, in this chapter we present a distributed FDI architecture specifically designed for uncertain networked nonlinear large-scale systems. We will consider different sources of uncertainty, namely modeling uncertainty, measurement noise, network-related uncertainties, such as communication delays, packet losses and asynchronous measurements, and the presence of possibly unknown anomalies. In Section 1.2 the problem formulation is given and the objectives and contributions of this chapter are explained in detail. In Section 1.3 the development of a fault detection scheme is presented in a continuous time framework based on [48], where a filtering technique, which is embedded in the design of the residual and threshold signals, is used to attenuate the measurement noise. This allows for the design of tight thresholds, and thus enhances fault detectability whilst guaranteeing the absence of false-alarms. This filtering approach for fault detection is rigorously investigated, providing results regarding the class of detectable faults, the magnitude of detectable faults and the filtering impact (according to the poles' location and filters' order) on the detection time.

Section 1.4 addresses the need for integration between the different levels composing CPS systems, which are deeply correlated in modern systems, by presenting a comprehensive architecture, based on [14], where all the parts of complex distributed systems are considered: the physical environment, the sensor level, the diagnosers layer and the communication networks. Based on the problem formulation given in Section 1.2 and on the filtering approach explained in Section 1.3, a distributed fault-diagnosis approach is designed for distributed uncertain nonlinear large-scale systems to specifically address the issues emerging when considering networked diagnosis systems, such as the presence of delays and packet dropouts in the communication networks that degrade performance and could be a source of instability, misdetection, and false alarms.

Section 1.5 discusses some issues regarding fault diagnosis, that is the actions taken after the detection of a fault, for identifying its location and its magnitude or even learning the fault function so that it can be used for fault accommodation schemes. Finally, in Section 1.6 some concluding remarks are given.

1.2 Problem Formulation

Consider a large-scale distributed nonlinear dynamic system composed of N subsystems Σ_I, $I \in \{1,...,N\}$, each of which is described by the differential equation:

$$\Sigma_I : \begin{cases} \dot{x}_I(t) = f_I(x_I(t),u_I(t)) + g_I(x_I(t),z_I(t),u_I(t)) + \eta_I(x_I(t),z_I(t),u_I(t)) \\ \qquad +\beta_I(t-T_0)\phi_I(x_I(t),z_I(t),u_I(t)) & (1.1) \\ m_I(t) = x_I(t) + w_I(t), & (1.2) \end{cases}$$

where $x_I \in \mathbb{R}^{n_I}$, $u_I \in \mathbb{R}^{l_I}$ and $m_I \in \mathbb{R}^{n_I}$ are the state, input and measured output vectors of the I-th subsystem respectively, $z_I \in \mathbb{R}^{\bar{n}_I}$ is the vector of interconnection variables which are the state variables of the other subsystems $J \in \{1,\ldots,N\} \setminus \{I\}$ that affect the I-th subsystem, $f_I : \mathbb{R}^{n_I} \times \mathbb{R}^{l_I} \mapsto \mathbb{R}^{n_I}$ is the known local function dynamics of the I-th subsystem and $g_I : \mathbb{R}^{n_I} \times \mathbb{R}^{\bar{n}_I} \times \mathbb{R}^{l_I} \mapsto \mathbb{R}^{n_I}$ is the known part of the interconnection function between the I-th and the other subsystems. The vector function $\eta_I : \mathbb{R}^{n_I} \times \mathbb{R}^{\bar{n}_I} \times \mathbb{R}^{l_I} \mapsto \mathbb{R}^{n_I}$ is the overall modeling uncertainty associated with the known local and interconnection function dynamics and $w_I \in \mathscr{D}_{w_I} \subset \mathbb{R}^{n_I}$ ($\mathscr{D}_{w_I}$ is a compact set) represents the measurement noise. The state vectors x_I, $I \in \{1,...,N\}$ are considered unknown whereas their noisy counterparts m_I are known. Analogously, in the case of the interconnection variable z_I, only its noisy counterpart $m_{zI}(t) = z_I(t) + \varsigma_I(t)$ is available, where $\varsigma_I(t)$ is composed by the components of w_J affecting the relevant components of m_J (as before J refers to a neighboring subsystem). The term $\beta_I(t-T_0)\phi_I(x_I,z_I,u_I)$ characterizes the fault function dynamics affecting the I-th subsystem including its time evolution. More specifically, the term $\phi_I : \mathbb{R}^{n_I} \times \mathbb{R}^{\bar{n}_I} \times \mathbb{R}^{l_I} \mapsto \mathbb{R}^{n_I}$ is the unknown fault function and the term $\beta_I(t-T_0) : \mathbb{R} \mapsto \mathbb{R}^+$ denotes the time evolution of the fault, where T_0 is the unknown time of the fault occurrence [70]. Note that the fault function ϕ_I may depend on the interconnection state variable vector z_I and not only on the local state vector x_I. In this work we consider the case of a single fault that occurs in a subsystem (hence there is only one function $\phi_I(\cdot)$) and not the case of a distributed fault that spans across several subsystems. Of course, the fault that occurs in a subsystem Σ_I can affect neighboring subsystems Σ_J through the interconnection terms z_J. The fault time profile $\beta_I(t-T_0)$ can be used to model abrupt faults or incipient faults using a decaying exponential type function:

$$\beta_I(t-T_0) \triangleq \begin{cases} 0 & \text{if } t < T_0 \\ 1 - e^{-b_I(t-T_0)} & \text{if } t \geq T_0 \end{cases} \tag{1.3}$$

where $b_I > 0$ is typically an unknown parameter which denotes the fault evolution rate. Abrupt faults correspond to the limit $b_I \to \infty$, in this case, the time profile $\beta_I(t-T_0)$ becomes a step function. In general, small values of b_I indicate slowly developing faults (incipient faults) whereas large values of b_I make the time profile $\beta_I(t-T_0)$ approach a step function (abrupt faults).

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. $z_I(t)$, contains at least one of the state variables of Σ_J, i.e. $x_J(t)$.

The notation $|\cdot|$ used in this chapter indicates the absolute value of a scalar function or the 2-norm in case of a vector. In addition, the notation $y(t) = H(s)\big[x(t)\big]$ (which is used extensively in the adaptive control literature) denotes the output $y(t)$ of a linear system represented by the transfer function $H(s)$ with $x(t)$ as input. In terms of more rigorous notation, let $h(t)$ be the impulse response associated with $H(s)$; i.e. $h(t) \triangleq \mathscr{L}^{-1}\left[H(s)\right]$, where $\mathscr{L}^{-1}$ is the inverse Laplace transform operator. Then $y(t) = H(s)\big[x(t)\big] = \int_0^t h(\tau)x(t-\tau)\,\mathrm{d}\tau$.

The following assumptions are used throughout the chapter:

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

Assumption 2 *The modeling uncertainty $\eta_I^{(i)}$ (i denotes the i-th component of η_I) in each subsystem is an unstructured and possibly unknown nonlinear function of x_I, z_I and u_I but uniformly bounded by a known positive function $\bar{\eta}_I^{(i)}$, i.e.,*

$$|\eta_I^{(i)}(x_I, z_I, u_I)| \leq \bar{\eta}_I^{(i)}(m_I, m_{zI}, u_I), \quad i = 1,2,\ldots,n_I \tag{1.4}$$

for all $t \geq 0$ and for all $(x_I, z_I, u_I) \in \mathscr{D}_I$, where $m_{zI} = z_I + \varsigma_I$ is the measurable noisy counterpart of z_I, $\varsigma_I \in \mathscr{D}_{\varsigma_I} \subset \mathbb{R}^{\bar{n}_I}$ and $\bar{\eta}_I^{(i)}(m_I, m_{zI}, u_I) \geq 0$ is a known bounding function in some region of interest $\mathscr{D}_I = \mathscr{D}_{x_I} \times \mathscr{D}_{z_I} \times \mathscr{D}_{u_I} \subset \mathbb{R}^{n_I} \times \mathbb{R}^{\bar{n}_I} \times \mathbb{R}^{l_I}$. The regions $\mathscr{D}_{\varsigma_I}$ and $\mathscr{D}_I$ are compact sets. □

Assumption 1 is required for well-posedness since here we do not address the control design and fault accommodation. Assumption 2 characterizes the class of modeling uncertainties being considered. In practice, the system can be modeled more accurately in certain regions of the state space. Therefore, the fact that the bound $\bar{\eta}_I$ is a function of m_I, m_{zI} and u_I provides more flexibility by allowing the designer to take into consideration any prior knowledge of the system. Moreover, the bound $\bar{\eta}_I$ is required in order to distinguish the effects between modeling uncertainty and faults. For example if the bound $\bar{\eta}_I$ is not set properly and it is too low so that (1.4) does not hold, then false alarms may occur. On the other hand, if the bound $\bar{\eta}_I$ is set too high, so that (1.4) holds, then this might lead to conservative detection thresholds which may never be crossed, leading to undetected faults. Therefore, the handling of the modeling uncertainty is a key design issue in fault diagnosis architectures, which creates a trade-off between false alarms and conservative fault detection. In Section 1.4.4 adaptive approximation methods will be used to learn the modeling uncertainty η_I and use the learned function in order to obtain even tighter detection thresholds and enhance fault detectability.

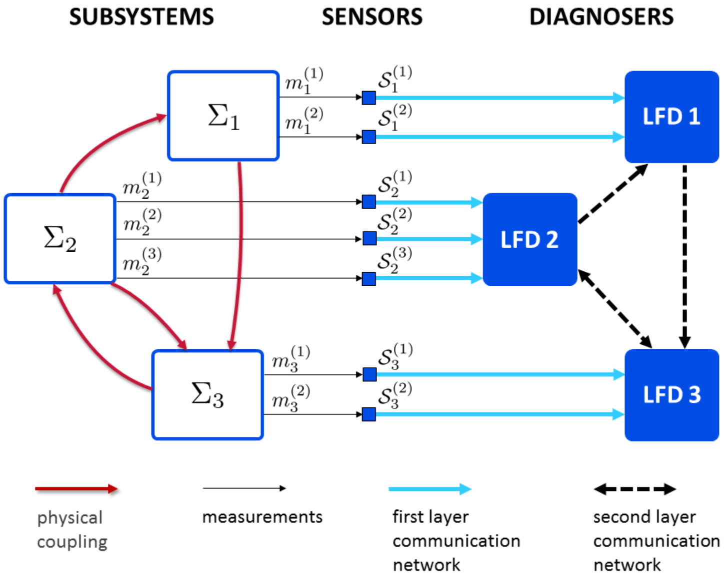

Fig. 1.2: An example of the proposed multi-layer fault detection architecture. The local state variables for each subsystem (physical layer, left) are measured by the sensor layer (center). The sensors communicate their measurements to the LFDs by means of the first level communication network. The second level communication network (right) allows the diagnosers to communicate with each other exchanging information.

Each sensor is associated with exactly one subsystem (see Fig. 1.2). The *local sensor* $S_I^{(i)}$ associated with the I-th subsystem provides a measurement $m_I^{(i)}$ of the i-th component of the local state vector x_I according to the output equation

$$S_I^{(i)} : m_I^{(i)}(t) = x_I^{(i)}(t) + w_I^{(i)}(t), \quad i = 1, \ldots, n_I, \tag{1.5}$$

where $w_I^{(i)}$ denotes the noise affecting the i-th sensor of the I-th subsystem.

Assumption 3 *For each i-th measurement $m_I^{(i)}$, with $i = 1, \ldots, n_I$, being the vector component index, the measurement uncertainty term $w_I^{(i)}$ is an unstructured and unknown function of time, but it is bounded by a known positive time-function $\bar{w}_I^{(i)}(t)$ such that $\left|w_I^{(i)}(t)\right| \leq \bar{w}_I^{(i)}(t)$, $i = 1, \ldots, n_I$, $I = 1, \ldots, N$, $t \geq 0$.* □

We assume that the control input is available without any error or delay (it is assumed that there exist feedback controllers yielding a local control action u_I such that some desired control objectives are achieved). Each subsystem is monitored by its respective Local Fault Diagnoser (LFD). The objective is to design and analyze a distributed fault detection scheme, with each subsystem Σ_I being monitored by a LFD that receives local measurements through the first communication network (see Fig. 1.2) and partial information (i.e. the measurements m_{zI} of the interconnection variables) from neighboring LFDs through the second communication network. In general, the distributed fault detection scheme is composed of N LFDs $\mathscr{S}_I$, one for each subsystem Σ_I. Each LFD $\mathscr{S}_I$ requires the input and output measurements of the subsystem Σ_I that it is monitoring and also the measurements of all interconnecting subsystems Σ_J that affect Σ_I. Note that these last measurements are communicated by neighboring LFDs $\mathscr{S}_J$, and not by the subsystems Σ_J. Therefore, there is the need of communication between the LFDs depending on their interconnections. It is important to note that, the second layer communication network mirrors the physical coupling morphology. Note that, the information exchanged among the subsystems is readily available since it is constituted by quantities z_I that are measurable with some uncertainty as $m_{zI}(t) = z_I(t) + \varsigma_I(t)$ (the noisy counterpart of z_I). Therefore, the distributed nature of the scheme stems from the fact that there is communication between the LFDs depending on their interconnections. More specifically, each LFD receives from its local sensors the noisy state measurements forming the vector $m_I = \mathrm{col}(m_I^{(i)}, i = 1,\dots,n_I)$ (see (1.5)) and, from the J-th neighboring LFD the noisy measurements $m_{zI}^{(i)}$, $i = 1,\dots,\bar{n}_I$ of the local state variables components $x_J^{(i)}$ that influence the I-th subsystem (i.e., the variables $x_J^{(i)}$ belonging to the interconnection vector z_I). Each LFD computes a local state estimate $\hat{x}_I(t)$ based on the local I-th model, by communicating the interconnection variables (and possibly other information) to neighboring LFDs. The LFD implements a model-based fault detection method: the local residual error vector $r_I(t)$ is compared, component–by–component, to a time-varying detection threshold vector $\bar{r}_I(t)$, suitably computed in order to guarantee the absence of false–alarms.

1.2.1 Objectives and Contributions

In this chapter, a distributed fault-diagnosis methodology is presented to address the sources of uncertainty mentioned in the introduction. More specifically:

a) a filtering-based design is embedded in a distributed fault-diagnosis methodology to dampen the effect of the measurement noise and enhance fault detection robustness by facilitating less conservative conditions for fault-detectability;

b) an adaptive learning approach is adopted to reduce the modeling uncertainty and thus, further enhance fault detectability;

c) a delay compensation strategy is devised to address delays and packet losses in the communication network between the LFDs using Time stamps and a buffer, called *diagnosis buffer* (see Fig.1.4);

d) a model–based re-synchronization algorithm is embedded in the diagnosis procedure to manage asynchronous measurements. This algorithm is based on *vir-*

tual sensors implemented in the LFDs and on the use of a *measurements buffer* (see Fig. 1.4);

In the following, we will first present in Section 1.3 the distributed filtering approach in a continuous-time framework under the assumptions of i) global synchronization, i.e. subsystems, sensors, and LFDs are assumed to share the same clock and sampling frequency, and ii) perfect information exchange, i.e., it is assumed that information exchanged between LFDs and communicated from the system to the LFDs is without any error nor delay and it is immediately available at any point of the diagnosis system. The effect of the filtering on the detectability performance is rigorously analyzed. After that, in Section 1.4, the filtering design is adapted in a discrete-time formulation to allow to analyze the more realistic networked scenarios, where different strategies for managing modeling uncertainty and network-related issues will be integrated in a comprehensive framework.

1.3 Filtering-based Distributed Fault Detection

In this section we present a filtering framework for the detection of faults in a class of interconnected, nonlinear, continuous-time systems with modeling uncertainty and measurement noise (see [48] for more details). In order to address the measurement noise issue which can lead to conservative detection thresholds or even false alarms if not dealt with properly, filtering is used by embedding the filters into the design in a way that takes advantage of the filtering noise suppression properties. Essentially, filtering dampens the effect of measurement noise in a certain frequency range allowing to set less conservative adaptive fault detection thresholds and thus enhancing fault detectability. As a result, a robust fault detection scheme is designed which guarantees no false alarms. The distributed fault detection scheme is comprised of a set of interacting LFDs, in which each subsystem is monitored by its respective detection agent.

To dampen the effect of measurement uncertainty $w_I(t)$, each measured variable $m_I^{(i)}$ is filtered by $H(s)$, where $H(s)$ is a p-th order filter with strictly proper transfer function

$$H(s) = sH_p(s), \tag{1.6}$$

$$H_p(s) = \frac{d_{p-2}s^{p-2} + d_{p-3}s^{p-3} + \ldots + d_0}{s^p + c_{p-1}s^{p-1} + \ldots + c_1 s + c_0}. \tag{1.7}$$

Note that the strictly proper requirement is important. If the transfer function $H(s)$ is proper, then the noise would appear in the filter output and the noise dampening would not be effective.

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.

Generally, each measured variable $m_I^{(i)}(t)$ can be filtered by a different filter. In this chapter, without loss of generality, we consider $H(s)$ to be the same for all the output variables in order to simplify the notation and presentation.

The filters $H(s)$ and $H_p(s)$ are asymptotically stable and hence BIBO stable. Therefore, for bounded measurement noise $w_I(t)$ (see Assumption 3), the filtered measurement noise $\varepsilon_{w_I}(t) \triangleq H(s)\,[w_I(t)]$ is uniformly bounded as follows:

$$|\varepsilon_{w_I}^{(i)}(t)| \leq \bar{\varepsilon}_{w_I}^{(i)} \quad i = 1, 2, \ldots, n_I, \tag{1.8}$$

where $\bar{\varepsilon}_{w_I}^{(i)}$ are known bounding constants. Depending on the noise characteristics, $H(s)$ can be selected to reduce the bound $\bar{\varepsilon}_{w_I}^{(i)}$.

1.3.1 Distributed Fault Detection

In this section, we explain in detail the fault filtering framework in order to obtain the residual signals $r_I(t)$ to be used for fault detection and the corresponding detection thresholds $\bar{r}_I(t)$. The fault detection logic is based on deriving suitable detection thresholds so that in the absence of a fault the residual signals are bounded by their corresponding detection threshold signals, guaranteeing no false alarms. To state this formally: in the absence of a fault (i.e. for $t \in [0, T_0)$), it is guaranteed that $|r_I^{(i)}(t)| \leq \bar{r}_I^{(i)}(t)$, $\forall i = 1, \ldots, n_I$ and $\forall I = 1, \ldots, N$. The detection decision of a fault in the overall system is made when $|r_I^{(i)}(t)| > \bar{r}_I^{(i)}(t)$ at some time t for at least one component i in any subsystem Σ_I. Note that, in this chapter, only a single fault ϕ_I is considered to occur in the large-scale distributed system.

By locally filtering the output signal $m_I(t)$ we obtain the filtered output $y_{I,f}(t)$:

$$\begin{aligned} y_{I,f}(t) &= H(s)\,[m_I(t)] \\ &= H(s)\,[x_I(t) + w_I(t)]\,. \end{aligned} \tag{1.9}$$

By using $\varepsilon_{w_I}(t) = H(s)\,[w_I(t)]$ and the fact that $s[x_I(t)] = \dot{x}_I(t) + x_I(0)\delta(t)$ (where $\delta(t)$ is the delta function), we obtain:

$$\begin{aligned} y_{I,f}(t) &= H(s)\,[x_I(t)] + \varepsilon_{w_I}(t) \\ &= H_p(s)\,[\dot{x}_I(t)] + H_p(s)\,[x_I(0)\delta(t)] + \varepsilon_{w_I}(t) \\ &= H_p(s)\big[f_I\big(x_I(t), u_I(t)\big) + g_I(x_I(t), z_I(t), u_I(t)\big) \\ &\qquad + \eta_I\big(x_I(t), z_I(t), u_I(t)\big) + \beta_I(t - T_0)\phi_I\big(x_I(t), z_I(t), u_I(t)\big)\big] \\ &\quad + \varepsilon_{w_I}(t) + h_p(t)x_I(0), \end{aligned} \tag{1.10}$$

where $h_p(t)$ is the impulse response of the filter $H_p(s)$, i.e. $h_p(t) \triangleq \mathscr{L}^{-1}[H_p(s)]$. The estimation model $\hat{x}_I(t)$ for $x_I(t)$ under fault-free operation is generated based on (1.1) by considering only the known components and by using the measurements m_I and m_{zI} as follows:

$$\dot{\hat{x}}_I = f_I(m_I(t), u_I(t)) + g_I(m_I(t), m_{zI}(t), u_I(t)), \tag{1.11}$$

with the initial condition $\hat{x}_I(0) = m_I(0)$.
The corresponding estimation model for $y_{I,f}(t)$, denoted by $\hat{y}_{I,f}(t)$, is given by

$$\hat{y}_{I,f}(t) = H(s)\big[\hat{x}_I(t)\big], \tag{1.12}$$

and by using (1.11) and following a similar procedure as in the derivation of (1.10), $\hat{y}_{I,f}(t)$ becomes:

$$\hat{y}_{I,f}(t) = H_p(s)\big[f_I\big(m_I(t), u_I(t)\big) + g_I\big(m_I(t), m_{zI}(t), u_I(t)\big)\big] + h_p(t)m_I(0). \tag{1.13}$$

The local residual error $r_I(t)$ to be used for fault detection is defined as:

$$r_I(t) \triangleq y_{I,f}(t) - \hat{y}_{I,f}(t), \tag{1.14}$$

and it is readily computable from equations (1.9), (1.11) and (1.12).

Prior to the fault ($t < T_0$), the local residual error can be written using equations (1.10), (1.13) and (1.14) as:

$$r_I(t) = H_p(s)\,[\chi_I(t)] + \varepsilon_{w_I}(t) \tag{1.15}$$

where the total uncertainty term $\chi_I(t)$ is defined as:

$$\chi_I(t) \triangleq \Delta f_I(t) + \Delta g_I(t) + \eta_I\big(x_I(t), z_I(t), u_I(t)\big), \tag{1.16}$$

$$\Delta f_I(t) \triangleq f_I\big(x_I(t), u_I(t)\big) - f_I\big(x_I(t) + w_I(t), u_I(t)\big), \tag{1.17}$$

$$\Delta g_I(t) \triangleq g_I\big(x_I(t), z_I(t), u_I(t)\big) - g_I\big(x_I(t) + w_I(t), z_I(t) + \varsigma_I(t), u_I(t)\big). \tag{1.18}$$

For simplicity, in the derivation of (1.15) the initial conditions $x_I(0) = m_I(0)$ are assumed to be known. If there is uncertainty in the initial conditions (i.e. $x_I(0) \neq m_I(0)$) then that introduces the extra term $h_p(t)(x_I(0) - m_I(0))$ in (1.15) which however converges to zero exponentially (since $h_p(t)$ is exponentially decaying [24]) and thus does not affect significantly the subsequent analysis.

By taking bounds on (1.15) and by using the triangle inequality for each component i of the residual, we obtain:

$$\begin{aligned}|r_I^{(i)}(t)| &\leq |H_p(s)\left[\chi_I^{(i)}(t)\right]| + |\varepsilon_{w_I}^{(i)}(t)| = |\int_0^t h_p(t-\tau)\chi_I^{(i)}(\tau)\,\mathrm{d}\tau| + |\varepsilon_{w_I}^{(i)}(t)| \\ &\leq \int_0^t |h_p(t-\tau)||\chi_I^{(i)}(\tau)|\,\mathrm{d}\tau + |\varepsilon_{w_I}^{(i)}(t)| \\ &\leq \int_0^t \bar{h}_p(t-\tau)\bar{\chi}_I^{(i)}(\tau)\,\mathrm{d}\tau + \bar{\varepsilon}_{w_I}^{(i)} \end{aligned} \tag{1.19}$$

where $\bar{h}_p(t)$ is the impulse response (of the filter $\bar{H}_p(s)$) that satisfies $|h_p(t)| \leq \bar{h}_p(t)$ for all $t > 0$ (details for selecting $\bar{H}_p(s)$ will be given in Section 1.3.2) and, $\bar{\chi}_I^{(i)}(t)$ is the bound on the total uncertainty term $\chi_I^{(i)}(t)$, i.e. $|\chi_I^{(i)}(t)| \leq \bar{\chi}_I^{(i)}(t)$.
Using Assumption 2, the bound $\bar{\chi}_I^{(i)}(t),\ i = 1,2,\ldots,n_I$ is defined as:

$$\bar{\chi}_I^{(i)}(t) \triangleq \overline{\Delta f}_I^{(i)} + \overline{\Delta g}_I^{(i)} + \bar{\eta}_I^{(i)}\big(m_I(t), m_{zI}(t), u_I(t)\big), \tag{1.20}$$

where

$$\overline{\Delta f}_I^{(i)} \triangleq \sup_{\substack{(x_I,u_I)\in\mathscr{D}_{x_I}\times\mathscr{D}_{u_I} \\ w_I\in\mathscr{D}_{w_I}}} |f_I^{(i)}\big(x_I,u_I\big) - f_I^{(i)}\big(x_I+w_I,u_I\big)| \tag{1.21}$$

$$\overline{\Delta g}_I^{(i)} \triangleq \sup_{\substack{(x_I,z_I,u_I)\in\mathscr{D}_I \\ (w_I,\varsigma_I)\in\mathscr{D}_{w_I}\times\mathscr{D}_{\varsigma_I}}} |g_I^{(i)}\big(x_I,z_I,u_I\big) - g_I^{(i)}\big(x_I+w_I,z_I+\varsigma_I,u_I\big)|. \tag{1.22}$$

Since the regions $\mathscr{D}_I$, $\mathscr{D}_{w_I}$ and $\mathscr{D}_{\varsigma_I}$ are compact sets, the suprema in (1.21) and (1.22) are finite. In addition, note that the bound $\bar{\chi}_I^{(i)}(t)$ in (1.20) depends on t because of the bounding function $\bar{\eta}_I^{(i)}$.

Finally, a suitable detection threshold $\bar{r}_I^{(i)}(t)$ can be selected as the right hand side of (1.19) which can be rewritten as:

$$\bar{r}_I^{(i)}(t) = \bar{H}_p(s)\left[\bar{\chi}_I^{(i)}(t)\right] + \bar{\varepsilon}_{w_I}^{(i)}. \tag{1.23}$$

A practical issue that requires consideration is the derivation of the bound $\bar{\chi}_I^{(i)}(t)$ given in (1.20). Specifically, the derivation of $\bar{\chi}_I^{(i)}(t)$ requires the bounds $\overline{\Delta f}_I^{(i)}$ and $\overline{\Delta g}_I^{(i)}$ on $\Delta f_I^{(i)}(t)$ and $\Delta g_I^{(i)}(t)$, respectively. One approach for deriving the bound $\overline{\Delta f}_I^{(i)}$ in (1.21) is to consider a local Lipschitz assumption:

$$|f_I^{(i)}(x_I,u_I) - f_I^{(i)}(x_I+w_I,u_I)| \leq L_{f_I^{(i)}}|w_I| \tag{1.24}$$

where $L_{f_I^{(i)}}$ is the Lipschitz constant for the function $f_I^{(i)}(x_I,u_I)$ with respect to x_I in the region $\mathscr{D}_{x_I}$. Therefore, if we have a bound w_I^M on the measurement noise, i.e. $|w_I(t)| \leq w_I^M \quad \forall t > 0$, then we can obtain a bound on $\Delta f_I^{(i)}(t)$. A similar approach can be followed for $\Delta g_I^{(i)}(t)$.

Another way of obtaining a less conservative bound than $\bar{\chi}_I^{(i)}$ and, therefore further enhance fault detectability, is by exploiting the use of filtering which can be proved beneficial for dampening the mismatch function $\Delta f_I(t) + \Delta g_I(t)$ which results due to the measurement noise. Among the various filters one can select, some may lead to less conservative detection thresholds. Therefore, a significantly less conservative detection threshold without the need for the Lipschitz constants can be obtained by observing that the residual (1.15) can be written as

$$r_I(t) = H_p(s)\left[\eta_I\big(x_I(t), z_I(t), u_I(t)\big)\right] + H_p(s)\left[\Delta f_I(t) + \Delta g_I(t)\right] + \varepsilon_{w_I}(t) \tag{1.25}$$

and by making the following assumption:

Assumption 4 *The filtered function mismatch term* $\varepsilon_{\Delta_I}(t) \triangleq H_p(s)\left[\Delta f_I(t) + \Delta g_I(t)\right]$ *is uniformly bounded as follows:*

$$|\varepsilon_{\Delta_I}^{(i)}(t)| \leq \bar{\varepsilon}_{\Delta_I}^{(i)} \quad i = 1, 2, \ldots, n_I, \tag{1.26}$$

where $\bar{\varepsilon}_{\Delta_I}^{(i)}$ *is a known bounding constant.* □

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

Therefore, the detection threshold becomes

$$\bar{r}_I^{(i)}(t) = \bar{H}_p(s)\left[\bar{\eta}_I^{(i)}\big(m_I(t), m_{zI}(t), u_I(t)\big)\right] + \bar{\varepsilon}_{\Delta_I}^{(i)} + \bar{\varepsilon}_{w_I}^{(i)}. \tag{1.27}$$

Figure 1.3 illustrates the I-th LFD which includes the implementation of the local filtered fault detection scheme for the I-th subsystem resulting from equations (1.9), (1.11), (1.12), (1.14) and (1.23).

1.3.2 Selection of filter $\bar{H}_p(s)$

Two methods for selecting a suitable transfer function $\bar{H}_p(s)$ with impulse response $\bar{h}_p(t)$ such that $|h_p(t)| \leq \bar{h}_p(t)$ for all $t \geq 0$ are illustrated.

In general though, note that if the impulse response $h_p(t)$ is non-negative, i.e. $h_p(t) \geq 0$, for all $t \geq 0$, then the calculation of $\bar{H}_p(s)$ can be omitted. In this case $H_p(s)$ can be used instead of $\bar{H}_p(s)$ in (1.23), as it can easily be seen from (1.19) since $|h_p(t-\tau)| = h_p(t-\tau)$. Necessary and sufficient conditions for non-negative impulse response for a specific class of filters are given in [60].

- *First method.*

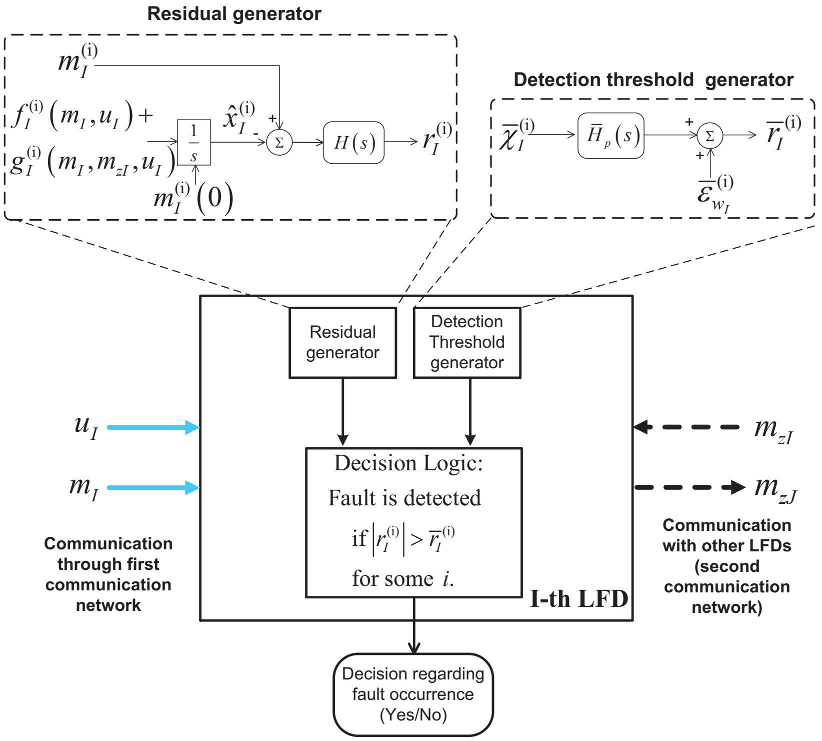
Residual generator
$m_I^{(i)}$
$f_I^{(i)}(m_I,u_I)+$
$g_I^{(i)}(m_I,m_{zI},u_I)$
$\frac{1}{s}$
$\hat{x}_I^{(i)}$
+
-
Σ
$H(s)$
$r_I^{(i)}$
$m_I^{(i)}(0)$
Detection threshold generator
$\bar{\chi}_I^{(i)}$
$\bar{H}_p(s)$
+
+
Σ
$\bar{r}_I^{(i)}$
$\bar{\varepsilon}_{w_I}^{(i)}$
Residual generator
Detection Threshold generator
Decision Logic:
Fault is detected
if $|r_I^{(i)}| > \bar{r}_I^{(i)}$
for some i.
I-th LFD
u_I
m_I
Communication through first communication network
m_{zI}
m_{zJ}
Communication with other LFDs (second communication network)
Decision regarding fault occurrence (Yes/No)

Fig. 1.3: Local Filtered Fault Detection Scheme

The first method relies on the following lemma, which describes a methodology for finding $\bar{H}_p(s)$. For notational convenience, for any $m \times n$ matrix A we define $|A|_{\mathscr{E}}$ as the matrix whose elements correspond to the modulus of the element $a_{i,j}$, $i = 1,\dots,m$ and $j = 1,\dots,n$ of the matrix A.

Lemma 1. ***[48].*** *Let $w(t) = Ce^{At}B$ be the impulse response of a strictly proper SISO transfer function $W(s)$ with state space representation (A,B,C). Then, for any signal $v(t) \geq 0$, the following inequality holds for all $t \geq 0$:*

$$\int_0^t |w(t-\tau)| v(\tau)\,\mathrm{d}\tau \leq \overline{W}(s)\,[v(t)]\,,$$

where $\overline{W}(s)$ is given by

$$\overline{W}(s) \triangleq |CT|_{\mathscr{E}}\,(sI - Re[J])^{-1}\,\left|T^{-1}B\right|_{\mathscr{E}} \tag{1.28}$$

and $J = T^{-1}AT$ is the Jordan form of the matrix A.

Therefore, by using Lemma 1 with $w(t) = h_p(t)$, the transfer function $\bar{H}_p(s)$ such that its impulse response satisfies $|h_p(t)| \leq \bar{h}_p(t)$ can be obtained from (1.28).

- *Second method.*

The second method is by using the following well-known result (see, for instance [24]).

Lemma 2. *The impulse response $h_p(t)$ of a strictly proper and asymptotically stable transfer function $H_p(s)$ decays exponentially; i.e., $|h_p(t)| \leq \kappa e^{-\upsilon t}$ for some $\kappa > 0$, $\upsilon > 0$, for all $t \geq 0$.*

By using Lemma 2, a suitable impulse response $\bar{h}_p(t)$ such that $|h_p(t)| \leq \bar{h}_p(t)$ for all $t \geq 0$ is given by $\bar{h}_p(t) = \kappa e^{-\upsilon t}$ and can be implemented using linear filtering techniques as $\bar{H}_p(s) = \frac{\kappa}{s+\upsilon}$.

1.3.3 Fault Detectability and Detection-Time Analysis

1.3.3.1 Fault Detectability Analysis

The design and analysis of the fault detection scheme in the previous sections were based on the derivation of suitable thresholds $\bar{r}_I^{(i)}(t)$ such that in the absence of any fault, the residual signals $r_I^{(i)}(t)$ are bounded by $\bar{r}_I^{(i)}(t)$. An important related question is what class of faults can be detected. This is referred to as *fault detectability analysis*. In this section, fault detectability conditions for the aforementioned fault detection scheme are derived. The fault detectability analysis constitutes a theoretical result that characterizes quantitatively the class of faults detectable by the proposed scheme.

Theorem 1. *Consider the nonlinear system (1.1), (1.2) with the distributed fault detection scheme described in (1.9), (1.11), (1.12), (1.14) and (1.23) in the general case of $H(s)$ given by (1.6). A sufficient condition for a fault $\phi_I^{(i)}(x_I, z_I, u_I)$ in the I-th subsystem initiated at T_0 to be detectable at time $T_d > T_0$ is that for some $i = 1, 2, \ldots, n_I$:*

$$\left|H_p(s)\left[\beta_I(T_d - T_0)\phi_I^{(i)}\left(x_I(T_d), z_I(T_d), u_I(T_d)\right)\right]\right| > 2\bar{r}_I^{(i)}(T_d). \tag{1.29}$$

Proof. In the presence of a fault that occurs at T_0, equation (1.15) becomes:

$$r_I^{(i)}(t) = H_p(s)\left[\chi_I^{(i)}(t) + \beta_I(t - T_0)\phi_I^{(i)}\left(x_I(t), z_I(t), u_I(t)\right)\right] + \varepsilon_{w_I}^{(i)}(t).$$

By using the triangle inequality, for $t > T_0$, the residual $r_I^{(i)}(t)$ satisfies:

$$\begin{aligned}
|r_I^{(i)}(t)| \geq & -|H_p(s)\big[\chi_I^{(i)}(t)\big]| - |\varepsilon_{w_I}^{(i)}(t)| \\
& + |H_p(s)\big[\beta_I(t-T_0)\phi_I^{(i)}\big(x_I(t), z_I(t), u_I(t)\big)\big]| \\
\geq & -\int_0^t |h_p(t-\tau)||\chi_I^{(i)}(\tau)|\,\mathrm{d}\tau - |\varepsilon_{w_I}^{(i)}(t)| \\
& + |H_p(s)\big[\beta_I(t-T_0)\phi_I^{(i)}\big(x_I(t), z_I(t), u_I(t)\big)\big]| \\
\geq & -\int_0^t \bar{h}_p(t-\tau)\bar{\chi}_I^{(i)}(\tau)\,\mathrm{d}\tau - \bar{\varepsilon}_{w_I}^{(i)} \\
& + |H_p(s)\big[\beta_I(t-T_0)\phi_I^{(i)}\big(x_I(t), z_I(t), u_I(t)\big)\big]| \\
\geq & -\bar{r}_I^{(i)}(t) + |H_p(s)\big[\beta_I(t-T_0)\phi_I^{(i)}(x_I(t), z_I(t), u_I(t))\big]|.
\end{aligned}$$

For fault detection, the inequality $|r_I^{(i)}(t)| > \bar{r}_I^{(i)}(t)$ must hold at some time $t = T_d$ for some $i = 1, \ldots, n_I$, so the final fault detectability condition given by (1.29) is obtained.

□

Although Theorem 1 is based on threshold (1.23), it can be readily shown that the same result holds in the case where threshold (1.27) is used. Clearly, the fault functions $\phi_I(x_I, z_I, u_I)$ are typically unknown and therefore this condition cannot be checked a priori. However, it provides useful intuition about the type of faults that are detectable. The detectability condition given in Theorem 1 is a sufficient condition, but not a necessary one and hence, the class of detectable faults can be significantly larger. The use of filtering is of crucial importance in order to derive tighter detection thresholds that guarantee no false alarms. As it can be seen in the detectability condition given by (1.29), the detection of the fault depends on the filtered fault function ϕ_I and as a result, the selection of the filter is very important. Since the fault function is usually comprised of lower frequency components, it is not affected that much by low-pass filtering in comparison to the measurement noise which is usually of higher frequency. In addition, filtering allows the derivation of tighter detection thresholds and, as a result, the fault detectability condition can be met more easily. Obviously, some filter selections may lead to less conservative thresholds than others.

The detectability properties of the proposed filtering approach are further investigated by considering a specific case for the filter $H_p(s)$:

$$H_p(s) = \frac{\alpha^p}{(s+\alpha)^p}. \tag{1.30}$$

This type of filter is well-suited for gaining further intuition since it contains two parameters p and α that denote the order of the filter and the pole location, respectively. More specifically, the order p of the filter regulates the damping effect of the high frequency noise, whereas the value α of the filter determines the cutoff frequency at which the damping begins. In general, more selective filter implementations can be made (i.e. Butterworth filters) which may have some implications in the filters required for the detection threshold implementation (due to the fact that the impulse response may not be always positive). But, the particular filter $H_p(s)$ given by (1.30) is perfectly suited for the investigation of the analytical properties of the filtering scheme. Note also that $H_p(s)$ has a non-negative impulse response $h_p(t)$ and therefore $\bar{H}_p(s)$ can be selected simply as $H_p(s)$.

In order to conduct this fault detectability analysis, we simplify Assumption 2 by considering a constant bounding condition. It is important to note that the constant bounding of the uncertainty may introduce additional conservativeness, thus reducing the advantage given by the tighter conditions obtained through the filtering.

Assumption 5 *The modeling uncertainty $\eta_I^{(i)}$ in each subsystem is an unstructured and possibly unknown nonlinear function of x_I, z_I and u_I but uniformly bounded by a known positive scalar $\bar{\eta}_I^{(i)}$, i.e.,*

$$|\eta_I^{(i)}(x_I, z_I, u_I)| \leq \bar{\eta}_I^{(i)}, \quad i = 1, 2, \ldots, n_I \tag{1.31}$$

for all $t \geq 0$ and for all $(x_I, z_I, u_I) \in \mathscr{D}_I$, where $\bar{\eta}_I^{(i)} \geq 0$ is a known bounding scalar in some region of interest $\mathscr{D}_I = \mathscr{D}_{x_I} \times \mathscr{D}_{z_I} \times \mathscr{D}_{u_I} \subset \mathbb{R}^{n_I} \times \mathbb{R}^{\bar{n}_I} \times \mathbb{R}^{l_I}$. □

By using the Lipschitz assumption stated in (1.24), along with the known constant bound w_I^M of the measurement uncertainty $|w_I|$ and, the constant bound on the modeling uncertainty $\bar{\eta}_I^{(i)}$, as stated in Assumption 5, the bound of the total uncertainty term $\bar{\chi}_I^{(i)}(t)$ takes a constant value $\bar{\chi}_I^{(i)}$. Then, Theorem 2, which follows, can be obtained (its proof can be found in [48]).

It must be pointed out that, although we use (1.23) for the detection threshold, the adaptation of the subsequent results in the case where the threshold is given by (1.27) is straightforward by simply replacing $\bar{\chi}_I^{(i)}$ with $\bar{\eta}_I^{(i)}$ and adding the term $\bar{\varepsilon}_{\Delta_I}^{(i)}$ along the term $\bar{\varepsilon}_{w_I}^{(i)}$ in what follows.

Theorem 2. *Consider the nonlinear system (1.1), (1.2) with the distributed fault detection scheme described in (1.9), (1.11), (1.12), (1.14) and (1.23) in the special case of $H_p(s)$ given by (1.30) and with $\bar{H}_p(s) = H_p(s)$. Suppose at least one component $\phi_I^{(i)}(x_I, z_I, u_I)$ of the fault vector $\phi_I(x_I, z_I, u_I)$ satisfies the condition*

$$|\phi_I^{(i)}(x_I(t'), z_I(t'), u_I(t'))| \geq M, \quad \forall\, t' \in [T_0, t], \tag{1.32}$$

for sufficiently large $t > T_0$ and is continuous in the time interval $t' \in [T_0, t]$. If $M > 2(\bar{\chi}_I^{(i)} + \bar{\varepsilon}_{w_I}^{(i)})$, then the fault will be detected, that is $|r_I^{(i)}(t)| > \bar{r}_I^{(i)}(t)$.

The aforementioned theorem is conceptually different from Theorem 1. More specifically, the detectability condition (1.29) of Theorem 1 allows the fault function $\phi_I^{(i)}$ to change sign. On the other hand, Theorem 2 states that if the fault function $\phi_I^{(i)}$ maintains the same sign over time and its magnitude is larger than $2(\bar{\chi}_I^{(i)} + \bar{\varepsilon}_{w_I}^{(i)})$ for sufficiently long, then the fault is guaranteed to be detected.

1.3.3.2 Detection Time Analysis

The detection time of a fault, that is, the time interval between the fault occurrence and its detection, plays a crucial role in fault diagnosis and it constitutes a form of performance criterion. When a fault is detected faster, then timely actions can be undertaken to avoid more serious or even disastrous consequences. It is worth noting that incipient faults are more difficult to detect, especially during their early stages, and as a result the detection time of an incipient fault is generally larger than that of an abrupt fault. In this section, an upper bound of the detection time is obtained in the case where a fault is detected according to Theorem 2. Moreover, we investigate the influence of the filter's order p and the pole location α on the upper bound of the detection time in order to derive some insight regarding the selection of p and α. The results are obtained for the general case of an incipient fault; concerning the dependence of the detection time on the filter's order p, only the abrupt fault case is addressed for the sake of simplicity.

Theorem 3. *Consider the nonlinear system (1.1), (1.2) with the distributed fault detection scheme described in (1.9), (1.11), (1.12), (1.14) and (1.23) in the special case of $H_p(s)$ given by (1.30) and with $\bar{H}_p(s) = H_p(s)$. If at least one component $\phi_I^{(i)}(x_I, z_I, u_I)$ of the fault vector $\phi_I(x_I, z_I, u_I)$ satisfies the condition*

$$\left|\phi_I^{(i)}\big(x_I(t'), z_I(t'), u_I(t')\big)\right| \geq M, \quad \forall\, t' \in [T_0, t] \tag{1.33}$$

where $M > 2(\bar{\chi}_I^{(i)} + \bar{\varepsilon}_{w_I}^{(i)})$ for sufficiently large $t > T_0$ and is continuous in the time interval $t' \in [T_0, t]$ such that the fault can be detected according to Theorem 2, then:

(a) A sufficient condition for fault detectability is given by:

$$q(t, T_0, \alpha) > \frac{2(p-1)!}{M}(\bar{\chi}_I^{(i)} + \bar{\varepsilon}_{w_I}^{(i)}). \tag{1.34}$$

where

$$q(t, T_0, \alpha) \triangleq q_1(t, T_0, \alpha) - q_2(t, T_0, \alpha) \tag{1.35}$$

$$q_1(t, T_0, \alpha) = \gamma\big(p, \alpha(t - T_0)\big), \tag{1.36}$$

$$q_2(t, T_0, \alpha) = \begin{cases} \frac{\alpha^p}{p}(t-T_0)^p e^{-\alpha(t-T_0)} & \text{if } \alpha = b_I \\ \frac{\alpha^p e^{-b_I(t-T_0)}}{(a-b_I)^p}\gamma\big(p, (\alpha - b_I)(t-T_0)\big) & \text{else,} \end{cases} \tag{1.37}$$

and $\gamma(\cdot)$ indicates the lower incomplete Gamma function, defined as $\gamma\big(p, z\big) \triangleq \int_0^z w^{p-1} e^{-w}\,\mathrm{d}w$.

(b) An upper bound on the detection time T_d of an incipient fault can be found by solving the equation:

$$q_1(T_d,T_0,\alpha)-q_2(T_d,T_0,\alpha)=\frac{2(p-1)!}{M}\bar{r}_I^{(i)}(T_d), \tag{1.38}$$

where $\bar{r}_I^{(i)}$ is given by

$$\bar{r}_I^{(i)}(t)=\frac{1}{(p-1)!}\bar{\chi}_I^{(i)}\gamma\big(p,\alpha t\big)+\bar{\varepsilon}_{w_I}^{(i)}. \tag{1.39}$$

(c) The upper bound T_d decreases monotonically as the value of α increases.
(d) In the case of abrupt faults, the upper bound on the detection time T_d increases as the order p of the filter increases.

The proof of Theorem 3 can be found in [48]. Part (b) of the above theorem establishes the mathematical equation whose solution gives an upper bound on the detection time. At this point, we must stress that, although we refer to the solution of the equation as the upper bound of the detection time (because of the requirement (1.32)), there are cases where the solution is the actual detection time. For instance, consider the case where the magnitude of the fault is $\big|\phi_I^{(i)}(x_I(t'),z_I(t'),u_I(t'))\big| = M, \quad \forall\, t' \in [T_0,t]$ and $M > 2(\bar{\chi}_I^{(i)}+\bar{\varepsilon}_{w_I}^{(i)})$. Then, the solution of (1.38) gives the actual detection time.

Part (c) of the theorem shows that by increasing the value of the pole α, the upper bound on the detection time (and sometimes the actual detection time as explained before) decreases. On the other hand, the value of α regulates the cutoff frequency of the filter where the damping begins, so the pole location has an inherent trade-off between noise damping and fault detection speed.

Part (d) of the theorem states that in the case of abrupt faults, the upper bound on the detection time increases as the order p of the filter increases. Although the proof is for the case of abrupt faults, the same behavior is observed in the case of incipient faults as well. An obvious downside of higher order filtering is the possible increased detection time. There is also a qualitative explanation for Part (d), as it has necessarily to do with the phase lag introduced by the filter which increases with p. Simply put, by increasing p results in increased phase lag or delay between the input and output signals of the filter and since the detectability of a fault relies on the filtered signals, the detection time increases according to the delay incurred.

Remark 1 *Prior to the occurrence of a fault, the residual differs from zero due to the effect of the filtered noise and filtered modeling uncertainty as indicated by (1.15). When a fault occurs, the residual is permanently contaminated by the filtered fault function as shown in the proof of Theorem 1. In general, the location of the poles simply affects the effectiveness of the noise dampening. To make things more clear, consider Theorems 2 and 3 which rely on the special case of the filter $H_p(s)$ given in (1.30). Theorem 2, states that in the case of a fault (abrupt or incipient), which satisfies the conditions given in the Theorem then the fault is guaranteed to be detected. Note that this is irrespective of the location of the filters' poles. In fact, as shown in Theorem 3, having faster poles results in a smaller upper bound on the detection time or even smaller actual detection time. In conclusion, the location of the poles does not limit the duration of the residual activation when a fault* *occurs, but instead the residual is permanently affected by the filtered fault function. Therefore, the location of the poles has an inherent trade-off between noise damping and fault detection speed.*

Simulation results showing the effectiveness of the illustrated techniques can be found in [48].

1.4 The Cyber-Physical Networked Architecture

In this section we present a cyber-physical networked fault detection architecture based on [14]. Let us note that the approach for distributed fault diagnosis of nonlinear uncertain large-scale systems that we have previously described is based on some underlying assumptions that may restrict its applicability, namely:

1. global synchronization: subsystems, sensors, and LFDs were assumed to share the same clock and sampling frequency;
2. perfect information exchange: it was assumed that information exchanged between LFDs and communicated from the system to the LFDs is without any error nor delay and it is immediately available at any point of the diagnosis system.

In several realistic contexts, 1) and 2) may not hold, and as a consequence, i) some faults may become undetectable due to the fact that LFDs make detection decisions based on outdated information; ii) delays in information exchange may cause longer detection times; iii) the lack of accurate and timely information may cause false-alarms.

In order to address these issues and the more complex nature of real CPS systems, we now consider a more comprehensive framework, where the previously proposed filtering design to reduce measurement noise is adapted in the current formulation in discrete time.

The proposed distributed fault-detection architecture is made of three layers: the system layer, the sensor layer and the diagnosis layer. In Fig. 1.2, this layout was shown in a pictorial way. These three layers are briefly described next.

The *system layer* refers to the large-scale system to be monitored. It is described by the continuous-time state equations for each subsystem Eq. (1.1) and the output equations (1.2).

The *sensor layer* consists of the available sensors taking measurements $m_I^{(i)}(t)$ in continuous-time (see (1.5)) and sampling and sending such measurements to the I-th LFD at time-instants $t_{sI}^{(i)}$ that are not necessarily equally-spaced in time. As we do not assume that the measurements delivered by the sensors are synchronized with each other, each measurement is labeled with a Time Stamp (TS) [94] to indicate the time instant $t_{sI}^{(i)}$ at which the measurements are taken by sensor $S_I^{(i)}$ in the time-coordinate t.

The communication between the sensors and the LFDs is achieved through the *first level communication network* (see Fig. 1.2). This network can introduce delays and packet losses, for instance because of collision between different sensors trying to communicate at the same time. Therefore, measurements communicated from the sensors to LFDs may be received at any time-instant.

The *Diagnosis layer* consists of the previously introduced LFDs providing a distributed fault-diagnosis procedure. The structure of each LFD is shown in Fig. 1.4. As previously mentioned, each LFD receives the measurements from specific sensors with the aim to provide local fault diagnosis decisions. The LFDs operate in a discrete-time synchronous time-frame $k \in \mathbb{Z}$ which turns out to be more convenient for handling any communications delays, as will be seen in the next sections. For the sake of simplicity, the sampling time of the discrete time-frame is assumed to be unitary and the reference time is common, that is, the origin of the discrete-time axis is the same as that of the continuous-time axis. Therefore, the operation of the LFDs is based on the local discrete-time models, which are the discrete-time version of local models (1.1):

$$x_I(k+1) = f_I(x_I(k), u_I(k)) + g_I(x_I(k), z_I(k), u_I(k)) + \eta_I(x_I(k), z_I(k), u_I(k)) \\ + \beta_I(k-k_0)\phi_I(x_I(k), z_I(k), u_I(k)), \quad (1.40)$$

where ϕ_I describes the local discretized fault effects, occurring at some discrete-time k_0 (that is, $\beta_I(k-k_0)\phi_I(x_I(k), z_I(k), u_I(k)) = 0, k < k_0$). Each LFD exchanges information with neighboring LFDs by means of the *second level communication network* (see right side of Fig.1.2 and Fig. 1.4). As we will see in the following, the exchanged information consists in the re-synchronized interconnection variables v_J. In Fig. 1.4, an example of a two LFDs architecture is presented to provide more insight into the structure of the proposed scheme.

In summary, two different and not reliable communication networks are considered in this work: the first level communication network allows each LFD to communicate with its local sensors and the second level communication network allows the communication between different LFDs for detection purposes. Both these communication networks may be subject to delays and packet losses. Given the different nature of the networks (the first is local, while the second is connecting different subsystems, which may be geographically apart), in the next section we provide two different strategies to manage communication issues: a re-synchronization method for the first level communication network and a delay compensation strategy for the second level communication network.

1.4.1 Re-synchronization at Diagnosis Level

Let us consider a state variable $x_I^{(i)}(t)$; as mentioned before, at time $t = t_{sI}^{(i)}$ the sensor $S_I^{(i)}$ takes the measurement $m_I^{(i)}(t_{sI}^{(i)})$ and sends it to the I-th LFD with a time-stamp

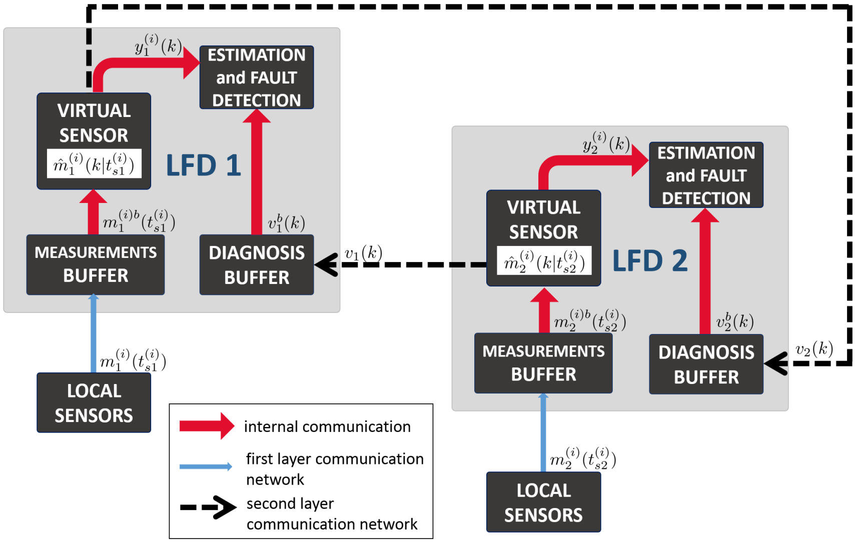
$y_1^{(i)}(k)$
ESTIMATION and FAULT DETECTION
VIRTUAL SENSOR
$\hat{m}_1^{(i)}(k|t_{s1}^{(i)})$
LFD 1
$m_1^{(i)b}(t_{s1}^{(i)})$
$v_1^b(k)$
MEASUREMENTS BUFFER
DIAGNOSIS BUFFER
$v_1(k)$
$m_1^{(i)}(t_{s1}^{(i)})$
LOCAL SENSORS
$y_2^{(i)}(k)$
ESTIMATION and FAULT DETECTION
VIRTUAL SENSOR
$\hat{m}_2^{(i)}(k|t_{s2}^{(i)})$
LFD 2
$m_2^{(i)b}(t_{s2}^{(i)})$
$v_2^b(k)$
MEASUREMENTS BUFFER
DIAGNOSIS BUFFER
$v_2(k)$
$m_2^{(i)}(t_{s2}^{(i)})$
LOCAL SENSORS
internal communication
first layer communication network
second layer communication network

Fig. 1.4: An example of a two LFDs architecture. The internal structure of each LFD is shown (similarly as in [14]), composed of two buffers (the measurements buffer and the diagnosis buffer) to collect the information received, respectively, by the local sensors and neighboring LFDs, the Virtual Sensor (processing the received measurements), and the Fault Detection unit, responsible for the monitoring analysis. The communicated information between LFDs is represented.

$t_{sI}^{(i)}$. The I-th diagnoser receives the measurement sent by $S_I^{(i)}$ at time $t_{aI}^{(i)} > t_{sI}^{(i)}$. Since the LFDs run the distributed fault-diagnosis algorithm with respect to a discrete-time framework associated with an integer k (see (1.40)), an on-line re-synchronization procedure has to be carried out at the Diagnosis level. Moreover, the possible time-varying delays and packet losses introduced by the communication networks between the local sensors and the corresponding LFDs have to be addressed since they may affect the fault diagnosis decision. Note that the classical discrete-time FD architecture assumes that quantities sampled at exactly time k are used to compute quantities related to time $k+1$. Unfortunately, the LFDs may receive measurements associated with time instants different from k, because of transmission delays and because of the arbitrary sampling time instants of the sensors. The availability of the time-stamp $t_{sI}^{(i)}$ enables each LFD to implement a set of *local virtual sensors* by which the re-synchronization of the measurements received at the Diagnosis level is implemented. We assume that sensors and diagnosers share the same clock at the local level[2].

[2] As example, this could be obtained in accordance with the IEEE 1588-2002 standard ("Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control

Specifically, each LFD collects the most recent sensors measurements in a buffer and computes a projection $\hat{m}_I^{(i)}(k|t_{sI}^{(i)})$ of these latest available measurements $m_I^{(i)}(t_{sI}^{(i)})$, $i = 1,\dots,n_I$, to the discrete time instant[3] $k \geq t_{aI}^{(i)} > t_{sI}^{(i)}$, by integrating the local nominal model on the time interval $[t_{sI}^{(i)}, k]$.

Remark 2 *Let us note that measurements may be related to and could be received also before time $k-1$, without any assumption on the delay length, thus allowing the possibility of measurement packet losses. Moreover, thanks to the use of the time stamps and the buffers, "'out-of-sequence"' packets can be managed. The same measurement could be used by the virtual sensor more than once to obtain more than one projections related to different discrete time instants.*

The projected measurement $\hat{m}_I^{(i)}(k|t_{sI}^{(i)})$ can be computed by noticing that, *under healthy mode of behavior*, the local nominal model (1.1) for the state component i at any time $t > t_{sI}^{(i)}$ can be rewritten as:

$$x_I^{(i)}(t) = x_I^{(i)}(t_{sI}^{(i)}) + \int_{t_{sI}^{(i)}}^{t} [f_I^{(i)}(x_I(\tau), u_I(\tau)) + g_I^{(i)}(x_I(\tau), z_I(\tau), u_I(\tau)) \\ + \eta_I^{(i)}(x_I(\tau), z_I(\tau), u_I(\tau))] d\tau \,.$$

Hence, the LFD implements a *virtual sensor* that generates an estimate of the measurement at discrete-time k given by

$$\hat{m}_I^{(i)}(k|t_{sI}^{(i)}) = m_I^{(i)}(t_{sI}^{(i)}) \\ + \int_{t_{sI}^{(i)}}^{k} [f_I^{(i)}(\hat{m}_I(\tau|t_{sI}^{(i)}), u_I(\tau)) + g_I^{(i)}(\hat{m}_I(\tau|t_{sI}^{(i)}), \hat{m}_{zI}(\tau|t_{sI}^{(i)}), u_I(\tau)) \\ + \hat{\eta}_I^{(i)}(\hat{m}_I(\tau|t_{sI}^{(i)}), \hat{m}_{zI}(\tau|t_{sI}^{(i)}), u_I(\tau))] d\tau \,, \quad (1.41)$$

where $\hat{\eta}_I$ characterizes an adaptive approximator designed to learn the unknown modeling uncertainty function η_I [27] and $\hat{m}_{zI}$ are the projections of the measured interconnection variables m_{zI}. An example enhancing the re-synchronization procedure for one LFD monitoring a subsystem with three state variables is illustrated in Fig. 1.5.

Remark 3 *It is worth noting that the discrete-time index $k \in \mathbb{Z}$ represents kind of a "virtual Time Stamp" (vTS) computed by the LFDs after the re-synchronization task and communicated in the second level communication network between LFDs. This will be exploited in Section 1.4.2.*

Systems"), where each diagnoser can be selected as a synchronization master for the sensors that communicate with it.

[3] Recall that the sampling time of the diagnosers is supposed to be unitary for simplicity.

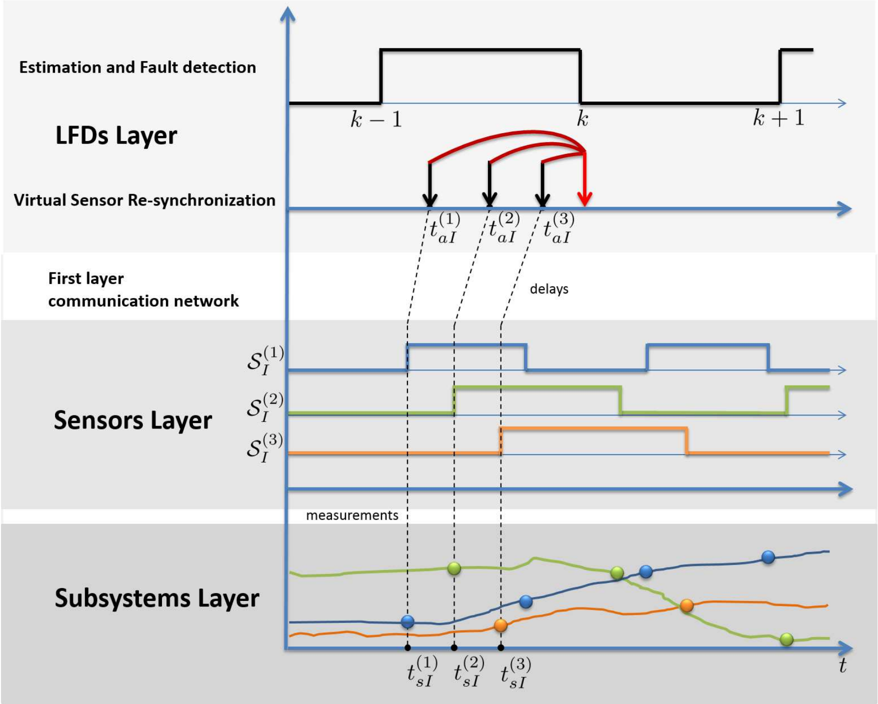

Fig. 1.5: The re-synchronization procedure [14] needed to manage delays and packet losses in the communication networks between each LFD and its local sensors. A single LFD is considered whose local model depends on three variables, which are measured by three different sensors. The clock signals of each layer involved are shown.

Remark 4 *Although in* (1.41)*, for analysis purposes,* $\hat{\eta}_I$ *represents the output of a continuous-time adaptive approximator, for implementation reasons, a suitable discrete-time approximator will be used, designed as explained in Section 1.4.4.*

The above-described projection and re-synchronization procedure gives rise to an additional source of measurement uncertainty: the *virtual measurement error*, which is defined as

$$\xi_I^{(i)}(k) \triangleq \hat{m}_I^{(i)}(k|t_{sI}^{(i)}) - x_I^{(i)}(k).$$

For the sake of analysis, it is worth noting that, due to synchronization and measurement noise, the virtual measurement error is given by:

$$\begin{aligned}\xi_I^{(i)}(k) =& m_I^{(i)}(t_{sI}^{(i)}) - x_I^{(i)}(t_{sI}^{(i)}) \\ &+ \int_{t_{sI}^{(i)}}^{k} [\Delta_{synch} f_I^{(i)}(\tau) + \Delta_{synch} g_I^{(i)}(\tau) + \Delta_{synch} \eta_I^{(i)}(\tau)] d\tau \\ =& w_I^{(i)}(t_{sI}^{(i)}) + \int_{t_{sI}^{(i)}}^{k} [\Delta_{synch} f_I^{(i)}(\tau) + \Delta_{synch} g_I^{(i)}(\tau) + \Delta_{synch} \eta_I^{(i)}(\tau)] d\tau \,, \end{aligned} \tag{1.42}$$

where

$$\Delta_{synch} f_I^{(i)}(\tau) \triangleq f_I^{(i)}(\hat{m}_I(\tau|t_{sI}^{(i)}), u_I(\tau)) - f_I^{(i)}(x_I(\tau), u_I(\tau)) \,,$$

$$\Delta_{synch} g_I^{(i)}(\tau) \triangleq g_I^{(i)}(\hat{m}_I(\tau|t_{sI}^{(i)}), \hat{m}_{zI}(\tau|t_{sI}^{(i)}), u_I(\tau)) - g_I^{(i)}(x_I(\tau), z_I(\tau), u_I(\tau)) \,,$$

and

$$\Delta_{synch} \eta_I^{(i)}(\tau) \triangleq \hat{\eta}_I^{(i)}(\hat{m}_I(\tau|t_{sI}^{(i)}), \hat{m}_{zI}(\tau|t_{sI}^{(i)}), u_I(\tau)) - \eta_I^{(i)}(x_I(\tau), z_I(\tau), u_I(\tau)) \,.$$

For notational convenience, we now collect the projected measurements $\hat{m}_I^{(i)}(k|t_{sI}^{(i)})$ in a vector, which, in the following, we denote as $y_I(k)$, with k being its vTS:

$$y_I(k) = \mathrm{col}\left\{ \hat{m}_I^{(i)}(k|t_{sI}^{(i)}), i = 1, \ldots, n_I \right\} .$$

Therefore, it is as if the virtual sensor implemented by the LFDs takes uncertain local measurements y_I of the state x_I, according to

$$y_I(k) = x_I(k) + \xi_I(k),$$

where ξ_I is the unknown virtual measurement error (1.42). Moreover, in place of the interconnection variables z_I, only the vector

$$v_I(k) = z_I(k) + \varsigma_I(k)$$

is available for diagnosis, as it is possible to see in Figure1.6, where ς_I is composed by the components of ξ_J affecting the relevant components of y_J (as before, J refers to a neighboring subsystem). For simplicity, we assume here that the control signal u_I is available to the diagnoser without any delays or other uncertainty.

The virtual measuring errors ξ_I and ς_I are unstructured and unknown. For each $i = 1, \ldots, n_I$ and $j = 1, \ldots, \bar{n}_I$, it is possible to compute a bound for their components using (1.42):

$$\left| \xi_I^{(i)}(k) \right| \leq \bar{\xi}_I^{(i)}(k), \qquad \left| \varsigma_I^{(j)}(k) \right| \leq \bar{\varsigma}_I^{(j)}(k),$$

where

$$\bar{\xi}_I^{(i)}(k) = \bar{w}_I^{(i)}(t_{sI}^{(i)}) + \int_{t_{sI}^{(i)}}^{k} \bar{\Delta}_{synch} f_I^{(i)}(\tau) + \bar{\Delta}_{synch} g_I^{(i)}(\tau) + \bar{\Delta}_{synch} \eta_I^{(i)}(\tau) d\tau \tag{1.43}$$

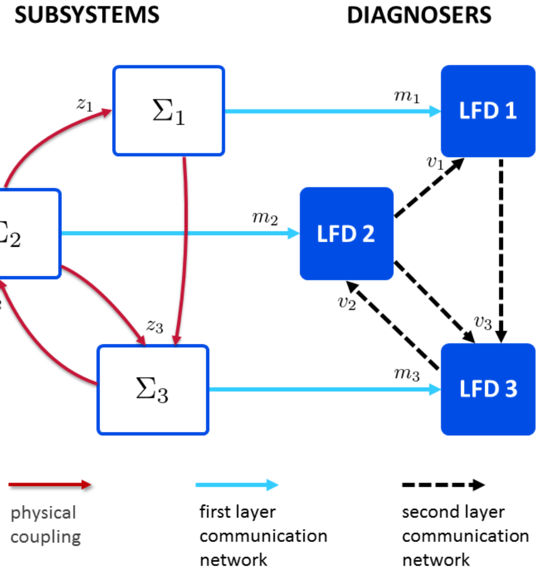

Fig. 1.6: An example of the multi-layer fault detection architecture. The interconnection variables z_I and the corresponding projected measurements v_I communicated among the diagnosers.

is a positive function, $\bar{w}_I^{(i)}$ is the one defined in Assumption 3,

$$\bar{\Delta}_{synch} f_I^{(i)}(\tau) = \max_{x_I \in \mathscr{R}^{n_I}} \left| f_I^{(i)}(\hat{m}_I(\tau), u_I(\tau)) - f_I^{(i)}(x_I(\tau), u_I(\tau)) \right|,$$

$$\bar{\Delta}_{synch} g_I^{(i)}(\tau) = \max_{x_I \in \mathscr{R}^{n_I}, z_I \in \mathscr{R}^{\bar{n}_I}} \left| g_I^{(i)}(\hat{m}_I(\tau), \hat{m}_{zI}(\tau), u_I(\tau)) - g_I^{(i)}(x_I(\tau), z_I(\tau), u_I(\tau)) \right|,$$

remembering that the sets $\mathscr{R}^{n_I}$, $\mathscr{R}^{\bar{n}_I}$ are the domain of the state and interconnection variables, respectively, and $\bar{\Delta}_{synch}\eta_I^{(i)}(\tau)$ can be computed in an analogous way as in (1.65) (see Section 1.4.6). The bound $\bar{\varsigma}_I$ is computed with the same procedure by the neighboring subsystems. In the next section, the fault-diagnosis procedure is presented.

1.4.2 The Distributed Fault Detection Methodology

For fault detection purposes, each LFD communicates with neighboring LFDs. It is assumed that the inter-LFD communication is carried over a packet-switched network, which we call the *second level communication network*, possibly subject to packet delays and losses. In order to manage delays in this network, the data-packets are Time Stamped, with the virtual Time Stamp, which contains the time instant the virtual measurements are referred to. In this layer, we assume to have perfect clock synchronization between the LFDs. In this way, all the devices of the monitoring architecture can share the same clock, that is, they know the reference time, and the use of Time Stamps can be valid.

Furthermore, we propose to provide each LFD with a buffer to collect the variables sent by neighbors. In the following, we denote with the superscript "b" the most recent value of a variable (or of a communicated function value) in the corresponding buffer of a given LFD; for example, v_I^b denotes the most recent value of the measured interconnection vector v_I contained in the buffer of the I-th LFD, while $[f_I(\cdot)]^b$ denotes the most recent value of the function $[f_I(\cdot)]$ in the buffer.

Each LFD computes a nonlinear adaptive estimate $\tilde{x}_I$ of the associated monitored subsystem state x_I. The local estimator, called *Fault Detection Approximation Estimator* (FDAE), is based on the local discrete-time nominal model (Eq.(1.40)). Similarly to what done in the first part of this chapter (Section 1.3), to dampen the effect of the virtual measurement error $\xi_I(k)$, each measured variable $y_I^{(i)} = x_I^{(i)} + \xi_I^{(i)}$ is filtered by $H(z)$, where $H(z)$ is a p-th order, asymptotically stable filter (poles lie inside the open unit disc $|z| = 1$) with proper transfer function

$$H(z) = \frac{d_0 + d_1 z^{-1} + d_2 z^{-2} + \ldots + d_p z^{-p}}{1 + c_1 z^{-1} + \ldots + c_p z^{-p}}. \tag{1.44}$$

Generally, each measured variable $y_I^{(i)}(k)$ can be filtered by a different filter but, without loss of generality, we consider $H(z)$ to be the same for all the output variables, in order to simplify notation and presentation. In addition, note that the form of $H(z)$ allows both IIR and FIR types of digital filters. The filter $H(z)$ can be written as $H(z) = zH_p(z)$ where $H_p(z)$ is the strictly proper transfer function

$$H_p(z) = \frac{d_0 z^{-1} + d_1 z^{-2} + d_2 z^{-3} + \ldots + d_p z^{-(p+1)}}{1 + c_1 z^{-1} + \ldots + c_p z^{-p}}. \tag{1.45}$$

Note that the filter $H_p(z)$ is also asymptotically stable since it comprises of the same poles as $H(z)$ with an additional pole at $z = 0$ (inside $|z| = 1$). Since the filters $H(z)$ and $H_p(z)$ (with impulse responses $h(t)$ and $h_p(t)$, respectively) are asymptotically stable, they are also BIBO stable. Therefore, for bounded virtual measurement error $\xi_I(k)$, the filtered virtual measurement error[4] $\Xi_I(k) \triangleq H(z)[\xi_I(k)]$ is bounded as follows:

$$\left|\Xi_I^{(i)}(k)\right| \leq \bar{\Xi}_I^{(i)}(k) \quad i = 1, \ldots, n_I \tag{1.46}$$

[4] For notational convenience, we use the shorthand $H(z)[\xi(k)]$ to denote $\mathscr{Z}^{-1}\{H(z)\Xi(z)\}$.

where $\bar{\Xi}_I^{(i)}$ are bounding functions that can be computed as $\bar{\Xi}_I^{(i)} \triangleq \bar{H}(z)[\bar{\xi}_I^{(i)}]$, being $\bar{H}(z)$ a filter with impulse response $\bar{h}(k)$ that satisfies $|h(k)| \leq \bar{h}(k)$ and using Eq. (1.43). The selection of suitable filters $\bar{H}(z)$ can be made by utilizing the methods indicated in Section 1.4.7. Note that we denote with capital letters the filtered signals.

1.4.3 Fault Detection Estimation and Residual Generation

In this subsection, we present a method for computing the local state estimate $\tilde{x}_I$ for fault detection purposes. The local estimation $\tilde{x}_I^{(i)}$ is given by

$$\begin{aligned}\tilde{x}_I^{(i)}(k+1) = f_I^{(i)}(y_I(k),u_I(k)) + g_I^{(i)}(y_I(k),v_I^b(k),u_I(k))\\ + \hat{\eta}_I^{(i)}(y_I(k),v_I^b(k),u_I(k),\hat{\vartheta}_I(k)),\end{aligned} \quad (1.47)$$

with initial condition $\tilde{x}_I^{(i)}(0) = y_I^{(i)}(0)$, where $\hat{\eta}_I$ is the output of an adaptive approximator designed in Subsection 1.4.4 to learn the unknown modeling uncertainty function η_I, $\hat{\vartheta}_I \in \hat{\Theta}_I$ denotes its adjustable parameters vector and t_b is the virtual time stamp of the most recent information received v_I^b in the buffer at time k.

The local estimation residual error $r_I(k)$ is defined as

$$r_I(k) \triangleq Y_I(k) - \widehat{Y}_I(k), \quad (1.48)$$

where we obtain the filtered output $Y_I(k)$ by locally filtering the measurement output signal $y_I(k)$

$$Y_I(k) \triangleq H(z)\left[y_I(k)\right], \quad (1.49)$$

and the output estimates as

$$\widehat{Y}_I(k) \triangleq H(z)\left[\tilde{x}_I(k)\right]. \quad (1.50)$$

The residual constitutes the basis of the fault detection scheme. It can be compared, component by component, to a suitable adaptive detection threshold $\bar{r}_I \in \mathbb{R}^{n_I}$, thus generating a local fault decision attesting the status of the subsystem: healthy or faulty. A fault in the overall system is said to be detected when $|r_I^{(i)}(k)| > \bar{r}_I^{(i)}(k)$, for at least one component i in any I-th LFD.

We now analyze the filtered measurements and estimates:

$$\begin{aligned}Y_I(k) &= H(z)\left[y_I(k)\right] = H(z)\left[x_I(k) + \xi_I(k)\right]\\ &= H_p(z)\left[z\left[x_I(k)\right]\right] + \Xi_I(k).\end{aligned} \quad (1.51)$$

In the absence of any faults (i.e., $\phi_I\big(x_I(k),z_I(k),u_I(k)\big) = 0$), (1.51) becomes

$$
\begin{aligned}
Y_I(k) &= H_p(z)\big[x_I(k+1) + z\big[x_I(0)\delta(k)\big]\big] + \Xi_I(k) \\
&= H_p(z)\big[f_I\big(x_I(k),u_I(k)\big) + g_I\big(x_I(k),z_I(k),u_I(k)\big) + \eta_I\big(x_I(k),z_I(k),u_I(k)\big)\big] \\
&\quad + h(k)x_I(0) + \Xi_I(k), \qquad (1.52)
\end{aligned}
$$

where $\delta(k)$ denotes the discrete-time unit-impulse sequence.

The filtered output estimation model for Y_I, denoted by $\widehat{Y}_I$, can be analyzed from the estimate provided by (1.47) as follows:

$$
\begin{aligned}
\widehat{Y}_I^{(i)}(k) =& H_p(z)\Bigg[f_I^{(i)}\big(y_I(k),u_I(k)\big) + g_I^{(i)}\big(y_I(k),v_I^b(k),u_I(k)\big) \\
&+ \hat{\eta}_I^{(i)}\big(y_I(k),v_I^b(k),u_I(k),\hat{\vartheta}_I(k)\big)\Bigg] + h(k)y_I^{(i)}(0). \qquad (1.53)
\end{aligned}
$$

Therefore, the residual (1.48) is readily computable from (1.49) and (1.50). The residual is analyzed in Subsection 1.4.6 to obtain a suitable adaptive detection threshold. Now, we design the adaptive approximator $\hat{\eta}_I$, needed to compute the state estimate (1.47) and hence (1.50).

1.4.4 Learning of the modeling uncertainty

Reducing the modeling uncertainty enables improved detection thresholds which, in turn, results in better detection capabilities. In this subsection, we consider the design of a nonlinear adaptive approximator, exploiting the variables available in the local buffers in each LFD to manage communication delays (the details of the delay compensation strategy are given in Subsection 1.4.5). The structure of the linear-in-the-parameters nonlinear multivariable approximator is not dealt with in this chapter (nonlinear approximation schemes like neural networks, fuzzy logic networks, wavelet networks, spline functions, polynomials, etc. can be used).

As shown later on in this subsection, adaptation of the parameters $\hat{\vartheta}_I$ of the approximator is achieved through the design of a dynamic state estimator which takes on the form:

$$
\hat{x}_I^{(i)}(k+1) = \lambda(\hat{x}_I^{(i)}(k) - y_I^{(i)}(k)) + f_I^{(i)}(y_I,u_I) + g_I^{(i)}(y_I,v_I^b,u_I) + \hat{\eta}_I^{(i)}(y_I,v_I^b,u_I,\hat{\vartheta}_I), \qquad (1.54)
$$

where $0 < \lambda < 1$ is a design parameter. Let us introduce the estimation error

$$
\varepsilon_I(k) \triangleq y_I(k) - \hat{x}_I(k)
$$

We compute the *i*-th state estimation error component as follows:

$$\begin{aligned}\varepsilon_I^{(i)}(k+1) &= y_I^{(i)}(k+1) - \hat{x}_I^{(i)}(k+1)\\ &= \lambda \varepsilon_I^{(i)} + \Delta f_I^{(i)} + \Delta g_I^{(i)} + \Delta \eta_I^{(i)} - \lambda \xi_I^{(i)} + \lambda \xi_I^{(i)}(k) + \xi_I^{(i)}(k+1)\,, \end{aligned} \quad (1.55)$$

where

$$\Delta f_I^{(i)} \triangleq f_I^{(i)}(x_I, u_I) - f_I^{(i)}(y_I, u_I)\,,$$

$$\Delta g_I^{(i)} \triangleq g_I^{(i)}(x_I, z_I, u_I) - g_I^{(i)}(y_I, v_I^b, u_I)\,,$$

and

$$\Delta \eta_I^{(i)} \triangleq \eta_I^{(i)}(x_I, z_I, u_I) - \hat{\eta}_I^{(i)}(y_I, v_I^b, u_I, \hat{\vartheta}_I)\,.$$

From this equation, the following learning law can be derived using Lyapunov stability techniques (see [107]) for every I:

$$\hat{\vartheta}_I(k+1) = P_{\hat{\Theta}_I}\left[\hat{\vartheta}_I(k) + \gamma_I L_I^\top [\varepsilon_I(k+1) - \lambda \varepsilon_I(k)\right]\,, \quad (1.56)$$

where $L_I^\top = \partial \hat{\eta}_I / \partial \hat{\vartheta}_I$ is the gradient matrix of the on-line approximator with respect to its adjustable parameters and $\gamma_I = \mu_I / \rho_I + \left\|L_I^\top\right\|_F^2$, with $P_{\hat{\Theta}_I}$ being a projection operator restricting $\hat{\vartheta}_I$ within $\hat{\Theta}_I$ [68], $\|\cdot\|_F$ denotes the Frobenius norm and $\rho_I > 0$, $0 < \mu_I < 2$ are design constants that guarantee the stability of the learning law [68].

1.4.5 Delay Compensation Strategy

Next, we analyze the properties of the Fault Detection estimator introduced in Subsection 1.4.3, where the filtered measurements are used; in particular, we explain how the estimator manages delays and packet losses in the second level communication network between diagnosers.

In order to compute (1.47) and (1.54), the generic J-th diagnoser communicates to the neighboring LFDs the current values of the variables v_I. It is worth noting that this information exchange between diagnosers can be affected by time-varying delays and packet losses and hence a compensation strategy has to be devised. The delay compensation strategy is derived without any assumption on the delay length, thus eventually dealing with the problem of packet losses and "out-of-sequence" packets. We assume that the communication network between diagnosers is designed so to avoid pathological scenarios, such as, for example, a situation in which the communication delay is always larger than the sampling time. It is important to note that a re-synchronization strategy like the one used in the first level communication networks cannot be used in this case, since here we consider data exchanged between different LFDs, and each LFD, of course, does not know the model of neighboring subsystems.

As in [12], thanks to the use of the virtual Time Stamps, the most recent measurements and information are considered. When a data packet arrives, its virtual Time Stamp v_{TS} is compared to t_b, which is the virtual Time Stamp of the information already in the buffer. If $v_{\mathrm{TS}} > t_b$, then the novel data packet takes its place in the buffer and $t_b \leftarrow v_{\mathrm{TS}}$. At time t_c, with $k < t_c < k+1$, each LFD computes the estimates for the time instant $k+1$ using information referred to time k. A variable in the buffer is up-to-date if $t_b = k$. Should a delay or a packet loss occur in the second level communication network, we proceed as follows. If some of the interconnection variables are not up-to-date, that is $t_b < k$, then the learning of the modeling uncertainty function η_I (1.56) is temporarily paused. Anyway, not up-to-date interconnection variables are used to compute the local value of the interconnection function in the state estimators (1.47) and (1.54), but this error is taken into account in the computation of the detection threshold, as will be seen in the following subsection.

1.4.6 Detection Threshold

In order to define an appropriate threshold for the detection of faults, we now analyze the dynamics of the output estimation error when the system is under healthy mode of behavior. Since, from (1.52) we have

$$\begin{aligned} Y_I^{(i)}(k) = H_p(z)\big[f_I^{(i)}\big(x_I(k),u_I(k)\big) + g_I^{(i)}\big(x_I(k),z_I(k),u_I(k)\big) \\ + \eta_I^{(i)}\big(x_I(k),z_I(k),u_I(k)\big)\big] + h(k)x_I^{(i)}(0) + \Xi_I^{(i)}(k), \end{aligned} \quad (1.57)$$

we are able to compute the residual defined in (1.48) by using (1.53) and (1.57):

$$r_I^{(i)}(k) = \Big[\chi_I^{(i)}(k)\Big]^b - \xi_I^{(i)}(0)h(k) + \Xi_I^{(i)}(k)\,, \quad (1.58)$$

where the total uncertainty term $\chi_I^{(i)}(k)$ is defined as:

$$\chi_I^{(i)}(k) \triangleq H_p(z)\big[\Delta f_I^{(i)}(k) + \Delta g_I^{(i)}(k) + \Delta \eta_I^{(i)}(k)\big]. \quad (1.59)$$

The function error $\Delta \eta_I$ can be computed as the sum of four different terms:

$$\Delta \eta_I = L_I \tilde{\vartheta}_I + \upsilon_I + \Delta \hat{\eta}_I + \Delta \eta_I^{\tau}. \quad (1.60)$$

The first term takes into account the error due to the parameters' estimation. This error can be characterized by introducing an *optimal weight vector* [98] $\hat{\vartheta}_I^*$ as follows:

$$\hat{\vartheta}_I^* \triangleq \arg\min_{\hat{\vartheta}_I} \sup_{x_I,z_I,u_I} \big\|\eta_I(x_I,z_I,u_I) - \hat{\eta}_I(x_I,z_I,u_I,\hat{\vartheta}_I)\big\|\,, \quad (1.61)$$

with $\hat{\vartheta}_I, x_I, z_I, u_I$ taking values in their respective domains, and by defining the parameter estimation error

$$\tilde{\vartheta}_I \triangleq \hat{\vartheta}_I^* - \hat{\vartheta}_I\,.$$

The second term in (1.60) is the so-called *Minimum Functional Approximation Error* υ_I, which describes the least possible approximation error that can be obtained at time k if $\hat{\vartheta}_I$ were optimally chosen:

$$\upsilon_I(k) \triangleq \eta_I(x_I, z_I, u_I) - \hat{\eta}_I(x_I, z_I, u_I, \hat{\vartheta}_I^*) \,.$$

Then, a term representing the error caused by the use of the uncertain measurements instead of the actual values of the state variables is defined:

$$\Delta\hat{\eta}_I \triangleq \hat{\eta}_I(x_I, z_I, u_I, \hat{\vartheta}_I) - \hat{\eta}_I(y_I, v_I, u_I, \hat{\vartheta}_I) \,.$$

Finally, the estimation error due to the use of delayed measurements is taken into account by

$$\Delta\eta_I^\tau \triangleq \hat{\eta}_I(y_I, v_I, u_I, \hat{\vartheta}_I) - \hat{\eta}_I(y_I, v_I^b, u_I, \hat{\vartheta}_I)$$

where v_I is the current measured variable and v_I^b is the value in the buffer, which is "old" in the presence of delays. Clearly, $\Delta\eta_I^\tau = 0$ when up-to-date measurements are used (in this case, $v_I^b = v_I$).

Using (1.60), the total uncertainty term $\chi_I^{(i)}(k)$ in (1.59) can be rewritten as

$$\begin{aligned}\chi_I^{(i)}(k) \triangleq H_p(z)\big[&\Delta f_I^{(i)}(k) + \Delta g_I^{(i)}(k) + L_I^{(i)}\tilde{\vartheta}_I(k) + \upsilon_I^{(i)}(k) \\ &+ \Delta\hat{\eta}_I^{(i)}(k) + \Delta\eta_I^{\tau(i)}(k)\big], \qquad (1.62)\end{aligned}$$

where $L_I^{(s_I)}$ indicates the s_I-th line of the matrix L_I. Using the triangle inequality, (1.58) satisfies:

$$\begin{aligned}\left|r_I^{(i)}(k)\right| &\le \left|\left[\chi_I^{(i)}(k)\right]^b\right| + \left|\xi_I^{(i)}(0)h(k)\right| + \left|\Xi_I^{(i)}(k)\right| \\ &\le \left[\left|\chi_I^{(i)}(k)\right|\right]^b + \bar{\xi}_I^{(i)}(0)\,|h(k)| + \bar{\Xi}_I^{(i)}(k). \qquad (1.63)\end{aligned}$$

From (1.62) and using again the triangle inequality, we obtain:

$$\begin{aligned}\left|\chi_I^{(i)}(k)\right| &\le \left|H_p(z)\big[\Delta f_I^{(i)}(k) + \Delta g_I^{(i)}(k) + \Delta\eta_I^{(i)}(k)\big]\right| \\ &\le \sum_{n=0}^{k} \left|h_p(k-n)\right| \Big|\Delta f_I^{(i)}(n) + \Delta g_I^{(i)}(n) + L_I^{(i)}\tilde{\vartheta}_I(n) + \upsilon_I^{(i)}(n) \\ &\quad + \Delta\hat{\eta}_I^{(i)}(n) + \Delta\eta_I^{\tau(i)}(n)\Big| \\ &\le \bar{\chi}_I^{(i)}(k) \triangleq \bar{H}_p(z)\big[\bar{\Delta} f_I^{(i)}(k) + \bar{\Delta} g_I^{(i)}(k) + \bar{\Delta}\eta_I^{(i)}(k)\big], \qquad (1.64)\end{aligned}$$

where $\bar{H}_p(z)$ is the transfer function with impulse response that satisfies $\left|h_p(k)\right| \le \bar{h}_p(k)$ (more details for the selection of $\bar{H}_p(z)$ are given in Subsection 1.4.7),

$$\bar{\Delta} f_I^{(i)}(k) \triangleq \max_{|\xi_I| \le \bar{\xi}_I} \left\{ \left| \Delta f_I^{(i)}(k) \right| \right\},$$

$$\bar{\Delta} g_I^{(i)}(k) \triangleq \max_{|\xi_I| \le \bar{\xi}_I(k)} \max_{|\varsigma_I| \le \bar{\varsigma}_I(k)} \left\{ \left| \Delta g_I^{(i)}(k) \right| \right\}$$

and

$$\begin{aligned} \bar{\Delta} \eta_I^{(i)}(k) \triangleq \left\| L_I^{(i)} \right\| \kappa_I(\hat{\vartheta}_I) + \bar{\upsilon}_I^{(i)}(k) + \max_{|\xi_I| \le \bar{\xi}_I(k)} \max_{|\varsigma_I| \le \bar{\varsigma}_I(k)} \left| \Delta \hat{\eta}_I^{(i)}(k) \right| \\ + \max_{v_I \in \mathscr{R}^v} \left| \hat{\eta}_I^{(i)}(y_I, v_I, u_I, \hat{\vartheta}_I) - \hat{\eta}_I^{(i)}(y_I, v_I^b(t_b), u_I, \hat{\vartheta}_I) \right|, \end{aligned} \quad (1.65)$$

with $\bar{\upsilon}_I$ denoting a bound to the minimum functional approximation error, the function κ_I being such that $\kappa_I(\hat{\vartheta}_I) \ge \left\| \tilde{\vartheta}_I \right\|$ and $\mathscr{R}^{v_I} \subset \mathbb{R}^{\bar{\eta}_I}$, where this last term represents a local domain of the interconnection variable and is communicated by the neighboring LFDs at $k = 0$. It is important to remark that $\mathscr{R}^{v_I}$ coincides with the domain $\mathscr{D}_{z_I}$ for subsystem I. Thanks to the way the threshold is designed from (1.63), it is straightforward that it guarantees the absence of false-alarms, since the residual *prior to the fault occurrence* always satisfies

$$\left| r_I^{(i)}(k) \right| \le \bar{r}_I^{(i)}(k),$$

where the detection threshold $\bar{r}_I^{(i)}$ is defined as

$$\bar{r}_I^{(i)}(k) \triangleq \left[\bar{\chi}_I^{(i)}(k) \right]^b + \bar{\xi}_I^{(i)}(0) \, |h(k)| + \bar{\Xi}_I^{(i)}(k). \quad (1.66)$$

Remark 5 *Notice that, even in the case of a conservative bound $\bar{\xi}_I^{(i)}$, the second term $\bar{\xi}_I^{(i)} \, |h(k)|$ affects the detection threshold only during the initial portion of the transient (the impulse response $h(k)$ of the filter $H(z)$ decays exponentially). Moreover, the term $\bar{\Xi}_I^{(i)}$ in* (1.65) *takes into account the uncertainty due to the delays in the communication network between LFDs. This term is instrumental to ensure the absence of false alarms caused by these communication delays.*

Remark 6 *The terms $\bar{\xi}_I(k)$ and $\bar{\varsigma}_I(k)$ are computed by the LFDs at each time-step after the re-synchronization task (see* (1.43)*) and are available to compute the fault detection threshold.*

Remark 7 *Admittedly, the bounds used in* (1.64) *and* (1.65) *give rise to conservative thresholds but have the advantage of guaranteeing the absence of false-positive alarms and of being easily computable requiring a small amount of data to be exchanged between the LFDs. In the presence of a-priori knowledge on the process to be monitored, tighter bound could be devised. For example, Lipschitz conditions on the local models could be easily exploited to devise tighter detection thresholds.*

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

A practical issue that requires consideration is the selection of the filter $\bar{H}_p(z)$ whose impulse response must satisfy $|h_p(t)| \leq \bar{h}_p(t)$ as stated before. In the case where the impulse response $h_p(t)$ is non-negative, the selection $\bar{H}_p(z) = H_p(z)$ is trivial. Sufficient conditions for non-negative impulse response for a class of discrete-time transfer functions are given in [60]. In the following, we present two methods for choosing $\bar{H}_p(z)$, 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. Due to the way $H_p(z)$ was defined, $H_p(z)$ is strictly proper and asymptotically stable. Hence, the impulse response $h_p(k)$ satisfies $|h_p(k)| \leq \kappa\lambda^k$ for all $k \in \mathbb{N}$, for some $\kappa > 0$ and $\lambda \in [0,1)$. Since $|h_p(k)| \leq \bar{h}_p(k)$ must hold, the impulse response $\bar{h}_p(k)$ can be selected as $\bar{h}_p(k) = \kappa\lambda^k$ and thus $\bar{H}_p(z) = \frac{\kappa}{1-\lambda z^{-1}}$.

Now, let's consider the case in which $H(z)$ is a FIR filter. FIR filters have several advantages, as they are inherently stable and can easily be designed to be linear-phase which corresponds to uniform delay at all frequencies. Let $H(z)$ be a p-th order FIR filter given by $H(z) = \sum_{n=0}^{p} d_n z^{-n}$. Therefore, $H_p(z) = z^{-1}H(z) = \sum_{n=0}^{p} d_n z^{-(n+1)}$ and $\bar{h}_p(k)$ can be selected as $\bar{h}_p(k) = |h_p(k)|$ which leads to the FIR filter $\bar{H}_p(z) = \sum_{n=0}^{p} |d_n| z^{-(n+1)}$.

1.4.8 The Local Fault Detection Algorithm

Now, all the elements needed to implement the fault detection scheme are available. For the sake of clarity, the implementation of the local fault detection methodology is sketched in the following Algorithm 1. Extensive simulation results showing the effectiveness of the presented approach can be found in [14].

1.4.9 Detectability Conditions

In this subsection, we address some sufficient conditions for detectability of faults by the proposed distributed networked fault detection scheme, thus considering the behavior of the fault detection algorithm in the case of a faulty system. We assume that at an unknown time k_0 a fault ϕ_I occurs. The fault detectability analysis constitutes a theoretical result that characterizes quantitatively (and implicitly) the class of faults detectable by the proposed scheme.

Theorem 4 (Fault Detectability). *A fault in the I-th subsystem occurring at time $k = k_0$ is detectable at a certain time $k = k_d$ if the fault function $\phi_I^{(i)}(x_I(k), z_I(k), u_I(k))$ satisfies the following inequality for some $i = 1, \ldots, n_I$:*

Algorithm 1 Fault detection algorithm for the *I*-th LFD

```
Learning = ON
Initialize the estimate $\hat{x}_I(0) = y_I(0)$
Initialize the estimate $\tilde{x}_I(0) = y_I(0)$
Compute the estimate $\hat{x}_I(1)$ (Eq. (1.54))
Compute the estimate $\tilde{x}_I(1)$ (Eq. (1.47))
Set $k = 1$
while A fault is not detected do
    Measurements $y_I(k)$ are acquired
    Compute $\varepsilon_I(k) = y_I(k) - \hat{x}_I(k)$ (for learning)
    Compute $Y_I(k)$ (Eq. (1.49)), $\widehat{Y}_I(k)$ (Eq. (1.50))
    Compute the residual $r_I(k) = Y_I(k) - \widehat{Y}_I(k)$
    Information from neighbors is acquired
    Compute the threshold $\bar{r}_I(k)$ (Eq. (1.66))
    Compare $|r_I(k)|$ with $\bar{r}_I(k)$
    if $|r_I(k)| > \bar{r}_I(k)$ then
        A fault is detected
        Learning = OFF
    end if
    if Some components $i$ of $v_I(k)$ are not received then
        Learning = OFF
    else
        Learning = ON
        $v_I^{b(i)}(k) = v_I^{(i)}(k)$
    end if
    if Learning = ON then
        Update $\hat{\vartheta}_I(k)$ (Eq. (1.56))
    else
        $\hat{\vartheta}_I(k) = \hat{\vartheta}_I(k-1)$
    end if
    Compute the novel estimate $\hat{x}_I(k+1)$ (Eq. (1.54))
    Compute the novel estimate $\tilde{x}_I(k+1)$ (Eq. (1.47))
    $k = k + 1$
end while
```

$$\left| \sum_{n=k_0}^{k_d} h_p(k-n)\phi_I^{(i)}\big(x_I(n), z_I(n), u_I(n)\big) \right| > 2\bar{r}_I^{(i)}(k_d). \tag{1.67}$$

Proof. After fault occurrence, that is for $k > k_0$, equation (1.58) becomes:

$$\begin{aligned} r_I^{(i)}(k) &= \chi_I^{(i)}(k)^b + H_p(z)\big[\phi_I^{(i)}\big(x_I(k), z_I(k), u_I(k)\big)\big] - \xi_I^{(i)}(0)h(k) + \Xi_I^{(i)}(k) \\ &= \chi_I^{(i)}(k)^b - \xi_I^{(i)}(0)h(k) + \Xi_I^{(i)}(k) + H_p(z)\big[\phi_I^{(i)}\big(x_I(k), z_I(k), u_I(k)\big)\big]. \end{aligned} \tag{1.68}$$

Using the triangle inequality, from (1.68) we can write

$$\begin{aligned}\left|r_I^{(i)}(k)\right| \geq &- \left|\chi_I^{(i)}(k)^b\right| - \left|\xi_I^{(i)}(0)h(k)\right| - \left|\Xi_I^{(i)}(k)\right| \\ &+ \left|H_p(z)\left[\phi_I^{(i)}\left(x_I(k), z_I(k), u_I(k)\right)\right]\right|\end{aligned} \tag{1.69}$$

and by using a similar procedure as in the derivation of (1.66), (1.69) becomes

$$\left|r_I^{(i)}(k)\right| \geq -\bar{r}_I^{(i)}(k) + \left|H_p(z)\left[\phi_I^{(i)}\left(x_I(k), z_I(k), u_I(k)\right)\right]\right|. \tag{1.70}$$

For fault detection at time $k = k_d$, the inequality $|r_I^{(i)}(k_d)| > \bar{r}_I^{(i)}(k_d)$ must hold for some $i = 1, \ldots, n_I$, so the final fault detectability condition is obtained:

$$\left|H_p(z)\left[\phi_I^{(i)}(x_I(k_d), z_I(k_d), u_I(k_d))\right]\right| > 2\bar{r}_I^{(i)}(k_d).$$

This can be rewritten in the summation form (1.67) of the Theorem.

□

This theorem provides a sufficient condition for the implicit characterization of a class of faults that can be detected by the proposed fault detection scheme. Let us note that the detectability condition represents the minimum cumulative magnitude of the fault that can be detected under a specific trajectory of the system. It is possible to study this condition off line for representative trajectories of the system.

1.4.10 Identification of the faulty subsystem

In the next section we consider the fault diagnosis problem. More specifically, we illustrate an approach for the adaptive learning of the local fault function after fault detection. Before developing the adaptive approximation procedure, we present an important remark.

A fundamental question regarding fault detectability is whether the fault that occurs in subsystem Σ_J is detectable not only by the LFD $\mathscr{F}_J$, but also by the LFD $\mathscr{F}_I$ of the neighboring subsystem Σ_I, whose state is influenced by Σ_J dynamics.

It can be shown (the interested reader can refer to [52]), that the proposed fault detection scheme guarantees that, a process fault $\phi_J(\cdot)$ occurring in subsystem Σ_J which affects Σ_I, can only be detected by its corresponding LFD $\mathscr{F}_J$ and not by the LFD $\mathscr{F}_I$. This result is essentially the implication of using the measurements of the state and interconnection variables in the estimation model given by (1.11). Qualitatively, this can be explained as follows. When a process fault 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 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 z_I and the detection LFD $\mathscr{F}_I$ makes use of the measurements v_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 LFD $\mathscr{F}_I$ and it is unable detect the fault. Hence, a process fault occurring in subsystem Σ_J is detectable only by its respective detection LFD $\mathscr{F}_J$ and not by any other LFD $\mathscr{F}_I$. This is a very important result because when a fault is detected in a subsystem, at the same time the faulty subsystem is identified, and further fault isolation/identification methods can be used targeting only the particular faulty subsystem.

1.5 Fault Diagnosis - Learning the Fault Function

After a fault is detected by the LFD $\mathscr{F}_I$ at time T_d, the fault isolation task is initiated to identify the type of fault occurring in the faulty subsystem Σ_I. In order to do this, various approaches can be used, and two of them are discussed in the sequel.

1.5.1 Generalised Observer Scheme

A fault isolation logic can be implemented based on a Generalized Observer Scheme (GOS, see [33, 65]). As in [31], it is assumed that each subsystem knows a *local fault set* $\mathscr{O}_I$, collecting all the $N_{\mathscr{O}_I}$ possible fault functions: $\phi_I^l(x_I, z_I, u_I)$, $l \in \{1, \dots, N_{\mathscr{O}_I}\}$. Once a fault is detected at time T_d in the I-th subsystem, the respective LFD $\mathscr{F}_I$ activates $N_{\mathscr{O}_I}$ estimators, where each filter is sensitive to a specific fault: the generic l–th fault isolation estimator of the I–th LFD is matched to the corresponding fault function ϕ_I^l, belonging to the local fault set $\mathscr{O}_I$. Each l–th estimator provides a local state estimate $\hat{x}_I^l$ of the local state x_I affected by the l-th fault:

$$\begin{aligned}\hat{x}_I^{l(i)}(k+1) = \lambda(\hat{x}_I^{l(i)}(k) - y_I^{(i)}(k)) + f_I^{(i)}(y_I, u_I) + g_I^{(i)}(y_I, v_I^b, u_I)\\ + \hat{\eta}_I^{(i)}(y_I, v_I^b, u_I, \hat{\vartheta}_I(T_d)) + \phi_I^{l(i)}(y_I, v_I^b, u_I),\end{aligned} \qquad (1.71)$$

where the learning of the modeling uncertainty has been stopped at time T_d in order not to learn the fault effect. The difference between the estimate $\hat{x}_I^l$ and the re-synchronized measurements y_I, after filtering, consists of the fault isolation estimation residual ${r_I}^l \triangleq Y_I - \hat{Y}_I^l$, where $\hat{Y}_I^l \triangleq H(z)[\hat{x}_I^l(k)]$. This residual is compared, component by component, to some properly designed isolation thresholds ${\bar{r}_I}^l$ so that if the j-th fault (in the fault set $\mathscr{O}_I$) has occurred, then it is guaranteed that

$$|{r_I}^{j(i)}(k)| \leq {\bar{r}_I}^{j(i)}(k) \quad \forall k > T_d, i = 1, \dots, n_I. \qquad (1.72)$$

The isolation thresholds are defined similarly as the detection threshold in (1.66), modifying $\bar{\chi}_I^{(i)}(k)$ adding the following term:

$$\bar{\Delta}\phi_I^{l(i)}(k) \triangleq \max_{|\xi_I| \le \bar{\xi}_I(k)} \max_{|\varsigma_I| \le \bar{\varsigma}_I(k)} \left\{ \left| \Delta\phi_I^{l(i)}(k) \right| \right\},$$

being $\Delta\phi_I^{l(i)}(k) = \phi_I^{l(i)}(x_I, z_I, u_I) - \phi_I^{l(i)}(y_I, v_I^b, u_I)$.

If a residual crosses its corresponding threshold, then we can exclude the occurrence of the considered l-th fault. Therefore, if we are able to exclude all the faults but one, then we can say that the fault is isolated.

1.5.2 Learning the fault function

In the case that the fault functions are not known a priori, we can use a different approach based on the adaptive learning of the fault function. According to the approximation model (1.54) introduced in Subsection 1.4.4 for learning the modeling uncertainty, when a fault is detected in the I-th subsystem, then the approximation model starts to learn the combined effect of the modeling uncertainty and the fault function. Assuming that the detection time T_d is sufficiently long, so that the modeling uncertainty is learned, its estimation is given by $\hat{\eta}_I(y_I(k), v_I^b(k), u_I(k), \hat{\vartheta}_I(T_d))$. Therefore, by allowing a sufficiently long learning period T_L after the fault detection, the approximator $\hat{\eta}_I$ learns the combined effect of the modeling uncertainty and the fault function as $\hat{\eta}_I(y_I(k), v_I^b(k), u_I(k), \hat{\vartheta}_I(T_d + T_L))$ for $k > T_d + T_L$. Therefore, the estimated fault function is given by $\hat{\phi}_I(k) = \hat{\eta}_I(y_I(k), v_I^b(k), u_I(k), \hat{\vartheta}_I(T_d + T_L)) - \hat{\eta}_I(y_I(k), v_I^b(k), u_I(k), \hat{\vartheta}_I(T_d))$, $k > T_d + T_L$. Note that, the fault could be incipient and still be developing at the end of the learning period, so the designer may let the learning process to continue. In this case, the estimated fault function is given by $\hat{\phi}_I(k) = \hat{\eta}_I(y_I(k), v_I^b(k), u_I(k), \hat{\vartheta}_I(k)) - \hat{\eta}_I(y_I(k), v_I^b(k), u_I(k), \hat{\vartheta}_I(T_d))$, $k > T_d + T_L$. The estimated fault function can then be used for fault accommodation purposes in order to guarantee the stability of the faulty system. For more information regarding this approach for learning the fault function, the interested reader can refer to [53].

1.6 Concluding Remarks

This chapter has reviewed a distributed fault diagnosis framework specifically designed for uncertain networked nonlinear large-scale systems concerning various sources of uncertainty, namely modeling uncertainty, measurement noise and network-related uncertainties.

In order to deal with the presence of measurement noise, a filtering scheme has been presented by integrating a general class of filters into the design of the residual and threshold signals in a way that takes advantage of the filtering noise suppression properties. Essentially, filtering dampens the effect of measurement noise in a certain frequency range allowing to set tighter detection thresholds and thus enhancing fault detectability. The main implications of the filtering scheme is rigorously investigated providing insights on the impact of the filters' poles and on the fault detection time.

The modeling uncertainties are also taken into account by means of an adaptive learning technique.

Furthermore, the chapter addressed the need for integration between the different levels composing CPS systems, by proposing a comprehensive architecture, where all parts of complex distributed systems are considered: the physical environment, the sensor level, the diagnosers layer and the communication networks. By adapting and incorporating the devised filtering scheme into the overall framework, a distributed fault-diagnosis approach has been designed for distributed uncertain nonlinear large-scale systems to specifically address the issues emerging when considering networked diagnosis systems, such as the presence of delays and packet dropouts in the communication networks that degrade performance and could be a source of instability, misdetection, and false alarms. Multi-rate systems, where the measurements may not be synchronous, were also considered. Under the stated assumptions, the proposed architecture guarantees the absence of false positive alarms.

Finally, some information was provided regarding the actions that can be taken after the detection of a fault in order to isolate the potential fault by identifying its location and magnitude, or even learning the fault function. Based on this information, actions can be taken in order to alleviate the fault effects and safeguard the system operation.

Modern, complex, interconnected systems can be prone to various sources of faults due to the increased complexity or even malicious attacks which can be considered as a "type" of fault. As a result, comprehensive fault diagnosis schemes need to be devised by considering the recent technological challenges, and this chapter has reviewed an integrated methodology which represents a step in that direction.

References

1. K. Adjallah, D. Maquin, and J. Ragot, "Nonlinear observer-based fault detection," in *IEEE Conference on Control Applications*, no. 3, 1994, pp. 1115–1120.
2. A. Ashari, R. Nikoukhah, and S. Campbell, "Active robust fault detection in closed-loop systems: Quadratic optimization approach," *IEEE Transactions on Automatic Control*, vol. 57, no. 10, pp. 2532–2544, 2012.
3. ——, "Effects of feedback on active fault detection," *Automatica*, vol. 48, no. 5, pp. 866–872, 2012.
4. K. Baheti and H. Gill, "Cyber-physical Systems," in *The Impact of Control Technology*, T. Samad and A. M. Annaswamy, Eds. IEEE Control Systems Society, 2011, pp. 161–166. [Online]. Available: http://ieeecss.org/general/impact-control-technology
5. R. Beard, "Failure accomodation in linear systems through self–reorganization," *Technical Report MTV-71-1, Man Vehicle Laboratory, MIT, Cambridge, MA*, 1971.
6. F. Blanchini, D. Casagrande, G. Giordano, S. Miani, S. Olaru, and V. Reppa, "Active fault isolation: A duality-based approach via convex programming," *SIAM Journal on Control and Optimization*, vol. 55, no. 3, pp. 1619–1640, 2017.
7. M. Blanke, M. Kinnaert, J. Lunze, and M. Staroswiecki, *Diagnosis and Fault Tolerant Control*. Berlin: Springer, 2003.

8. ——, “Distributed fault diagnosis and fault-tolerant control,” in *Diagnosis and Fault-Tolerant Control*. Springer, 2016, pp. 467–518.
9. ——, *Diagnosis and Fault-Tolerant Control*, 2nd ed. Springer Verlag, 2010.
10. S. Bodenburg and J. Lunze, “Plug-and-play reconfiguration of locally interconnected systems with limited model information,” *IFAC-PapersOnLine*, vol. 48, no. 22, pp. 20–27, 2015.
11. F. Boem, R. Carli, M. Farina, G. Ferrari-Trecate, and T. Parisini, “Scalable monitoring of interconnected stochastic systems,” in *2016 IEEE 55th Conference on Decision and Control (CDC)*, 2016, pp. 1285–1290.
12. F. Boem, R. Ferrari, T. Parisini, and M. Polycarpou, “Distributed fault detection for uncertain nonlinear systems: a network delay compensation strategy,” in *Proc. 2013 American Control Conference*, 2013.
13. F. Boem, S. Riverso, G. Ferrari-Trecate, and T. Parisini, “Plug-and-play fault detection and isolation for large-scale nonlinear systems with stochastic uncertainties,” *IEEE Transactions on Automatic Control (In press)*, 2018.
14. F. Boem, R. M. Ferrari, C. Keliris, T. Parisini, and M. M. Polycarpou, “A distributed networked approach for fault detection of large-scale systems,” *IEEE Transactions on Automatic Control*, vol. 62, no. 1, pp. 18–33, 2017.
15. F. Boem, R. M. Ferrari, and T. Parisini, “Distributed Fault Detection and Isolation of Continuous-Time Nonlinear Systems,” *European Journal of Control*, vol. 5-6, pp. 603–620, 2011.
16. F. Boem, R. M. Ferrari, T. Parisini, and M. M. Polycarpou, “Distributed fault diagnosis for continuous-time nonlinear systems: The input–output case,” *Annual Reviews in Control*, vol. 37, no. 1, pp. 163–169, 2013.
17. A. A. Cardenas, S. Amin, Z.-S. Lin, Y.-L. Huang, C.-Y. Huang, and S. Sastry, “Attacks against process control systems: risk assessment, detection, and response,” in *Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security*, ser. ASIACCS ’11. New York, NY, USA: ACM, 2011, pp. 355–366.
18. J. Chen and R. J. Patton, *Robust Model-Based Fault Diagnosis for Dynamic Systems*. Kluwer Academic Publishers Norwell, MA, USA, 1999.
19. P. Cheng, L. Shi, and B. Sinopoli, “Guest editorial special issue on secure control of cyber-physical systems,” *IEEE Transactions on Control of Network Systems*, vol. 4, no. 1, pp. 1–3, 2017.
20. S. Cheong and I. Manchester, “Input design for discrimination between classes of lti models,” *Automatica*, vol. 53, pp. 103–110, 2015.
21. L. Chiang, E. Russell, and R. Braatz, *Fault Detection and Diagnosis in Industrial Systems*. Springer-Verlag, London, 2001.
22. R. Clark, “Instrument fault detection,” *IEEE Transactions on Aerospace and Electronic Systems*, no. 3, pp. 456–465, 1978.
23. M. Davoodi, N. Meskin, and K. Khorasani, “Simultaneous fault detection and consensus control design for a network of multi-agent systems,” *Automatica*, vol. 66, pp. 185–194, 2016.
24. C. Desoer and M. Vidyasagar, *Feedback Systems: Input-Output Properties*, 1st ed. Academic Press, 1975.
25. P. Dorato, R. Tempo, and G. Muscato, “Bibliography on robust control,” *Automatica*, vol. 29, no. 1, pp. 201–213, 1993.
26. F. Dorfler, F. Pasqualetti, and F. Bullo, “Distributed detection of cyber-physical attacks in power networks: A waveform relaxation approach,” in *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*, sept. 2011, pp. 1486 –1491.
27. J. Farrell and M. M. Polycarpou, *Adaptive Approximation Based Control: Unifying Neural, Fuzzy, and Traditional Adaptive Approximation Approaches*. Hoboken, NJ: Wiley-Interscience, 2006.
28. J. Farrell, T. Berger, and B. Appleby, “Using learning techniques to accommodate unanticipated faults,” vol. 13, pp. 40—49, 1993.

29. H. Ferdowsi, D. Raja, and S. Jagannathan, “A decentralized fault prognosis scheme for nonlinear interconnected discrete-time systems,” in *American Control Conference*, 2012, pp. 5900–5905.
30. L. Ferranti, Y. Wan, and T. Keviczky, “Predictive flight control with active diagnosis and reconfiguration for actuator jamming.” *IFAC-PapersOnLine*, vol. 48, no. 23, pp. 166–171, 2015.
31. R. M. Ferrari, T. Parisini, and M. M. Polycarpou, “Distributed fault detection and isolation of large-scale nonlinear systems: an adaptive approximation approach,” *IEEE Transactions on Automatic Control*, vol. 57, no. 2, pp. 275–290, 2012.
32. E. Franco, R. Olfati-Saber, T. Parisini, and M. M. Polycarpou, “Distributed fault diagnosis using sensor networks and consensus-based filters,” in *Decision and Control, 2006 45th IEEE Conference on*. IEEE, 2006, pp. 386–391.
33. P. Frank, “Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy: A survey and some new results,” *Automatica*, vol. 26, no. 3, pp. 459–474, 1990.
34. P. Frank and S. X. Ding, “Survey of robust residual generation and evaluation methods in observer-based fault detection systems,” *Journal of Process Control*, no. 6, pp. 403–424, 1997.
35. Z. Gao, C. Cecati, and S. Ding, “A survey of fault diagnosis and fault-tolerant techniques—part i: Fault diagnosis with model-based and signal-based approaches,” *IEEE Transactions on Industrial Electronics*, vol. 62, no. 6, pp. 3757–3767, 2015.
36. E. A. Garcia and P. Frank, “Deterministic nonlinear observer-based approaches to fault diagnosis: a survey,” *Control Engineering Practice*, vol. 5, no. 5, pp. 663–670, 1997.
37. J. Gertler, “Survey of model-based failure detection and isolation in complex plants,” *IEEE Control Systems Magazine*, vol. 8, no. 6, pp. 3–11, 1988.
38. ——, “Fault detection and isolation using parity relations,” *Control Engineering Practice*, vol. 5, no. 5, pp. 653–661, May 1997.
39. ——, *Fault detection and diagnosis in engineering systems*, 1st ed. CRC Press, 1998.
40. V. Gupta and V. Puig, “Distributed fault diagnosis using minimal structurally overdetermined sets: Application to a water distribution network,” in *3rd Conference on Control and Fault-Tolerant Systems (SysTol)*. IEEE, 2016, pp. 811–818.
41. H. Hammouri, M. Kinnaert, and E. El Yaagoubi, “Observer-based approach to fault detection and isolation for nonlinear systems,” *IEEE Transactions on Automatic Control*, vol. 44, no. 10, pp. 1879–1884, 1999.
42. F. Harirchi, S. Yong, E. Jacobsen, and N. Ozay, “Active model discrimination with applications to fraud detection in smart buildings,” in *IFAC World Congress, Toulouse, France*, 2017.
43. I. Hwang, S. Kim, Y. Kim, and C. E. Seah, “A Survey of Fault Detection, Isolation, and Reconfiguration Methods,” *IEEE Transactions on Control Systems Technology*, vol. 18, no. 3, pp. 636–653, May 2010.
44. R. Isermann, *Fault-Diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance*. Springer-Verlag, 2006.
45. ——, “Process fault detection based on modeling and estimation methods - A survey,” *Automatica*, vol. 20, no. 4, pp. 387–404, Jul. 1984.
46. K. H. Johansson, G. J. Pappas, P. Tabuada, and C. J. Tomlin, “Guest editorial special issue on control of cyber-physical systems,” *IEEE Transactions on Automatic Control*, vol. 59, no. 12, pp. 3120–3121, 2014.
47. H. Jones, “Failure detection in linear systems,” Ph.D. Thesis, Dept. of Aero and Astro, MIT, Cambridge, MA, 1973.
48. C. Keliris, M. M. Polycarpou, and T. Parisini, “A Distributed Fault Detection Filtering Approach for a Class of Interconnected Continuous-Time Nonlinear Systems,” *IEEE Transactions on Automatic Control*, vol. 58, no. 8, pp. 2032–2047, 2013.
49. ——, “A Distributed Fault Detection Filtering Approach for a Class of Interconnected Input-Output Nonlinear Systems,” in *Proc. of European Control Conference*, 2013, pp. 422–427.

50. ——, “A Distributed Fault Diagnosis Approach Utilizing Adaptive Approximation for a Class of Interconnected Continuous-Time Nonlinear Systems,” in *Proc. of Control and Decision Conference*, 2014, pp. 6536–6541.
51. ——, “A Robust Nonlinear Observer-based Approach for Distributed Fault Detection of Input-Output Interconnected Systems,” *Automatica*, vol. 53, no. 3, pp. 408–415, 2015.
52. ——, “Distributed Fault Diagnosis for Process and Sensor Faults in a Class of Interconnected Input-Output Nonlinear Discrete-Time Systems,” *International Journal of Control*, 2015.
53. ——, “An Integrated Learning and Filtering Approach for Fault Diagnosis of a Class of Nonlinear Dynamical Systems,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 4, pp. 988–1004, Apr 2017.
54. S. Klinkhieo and R. J. Patton, “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*, 2009, pp. 1043–1048.
55. J. Lan and R. Patton, “Decentralized fault estimation and fault-tolerant control for large-scale interconnected systems: An integrated design approach,” in *UKACC 11th International Conference on Control*. IEEE, 2016, pp. 1–6.
56. N. Léchevin and C. Rabbath, “Decentralized Detection of a Class of Non-Abrupt Faults With Application to Formations of Unmanned Airships,” *IEEE Transactions on Control Systems Technology*, vol. 17, no. 2, pp. 484–493, 2009.
57. E. Lee, “Cyber physical systems: Design challenges,” in *Object Oriented Real-Time Distributed Computing (ISORC), 2008 11th IEEE International Symposium on*, may 2008, pp. 363 –369.
58. E. A. Lee, “Cyber-physical systems - are computing foundations adequate?” in *Position Paper for NSF Workshop On Cyber-Physical Systems: Research Motivation, Techniques and Roadmap*, October 2006.
59. J. Lee, B. Bagheri, and H.-A. Kao, “A cyber-physical systems architecture for industry 4.0-based manufacturing systems,” *Manufacturing Letters*, vol. 3, pp. 18–23, 2015.
60. Y. Liu and P. Bauer, “Sufficient conditions for non-negative impulse response of arbitrary-order systems,” in *IEEE Asia Pacific Conference on Circuits and Systems*, 2008, pp. 1410–1413.
61. M. W. Maier, “Architecting principles for systems-of-systems,” *Systems Engineering*, vol. 1, no. 4, pp. 267–284, 1998.
62. G. Marseglia and D. Raimondo, “Active fault diagnosis: A multi-parametric approach,” *Automatica*, vol. 79, pp. 223–230, 2017.
63. F. Pasqualetti, F. Dörfler, and F. Bullo, “Attack detection and identification in cyber-physical systems – part i: Models and fundamental limitations,” *ArXiv e-prints*, Feb. 2012.
64. ——, “Attack detection and identification in cyber-physical systems – part ii: Centralized and distributed monitor design,” *ArXiv e-prints*, Feb. 2012.
65. R. Patton, P. Frank, and D. Clark, *Fault Diagnosis in Dynamic Systems: Theory and Application*. Upper Saddle River, NJ, USA: Prentice Hall, 1989.
66. R. J. Patton, C. Kambhampati, A. Casavola, P. Zhang, S. X. Ding, and D. Sauter, “A generic strategy for fault-tolerance in control systems distributed over a network,” *European Journal of Control*, vol. 13, no. 2-3, pp. 280–296, 2007.
67. I. R. Petersen and R. Tempo, “Robust control of uncertain systems: Classical results and recent developments,” *Automatica*, vol. 50, no. 5, pp. 1315–1335, 2014.
68. M. M. Polycarpou, “On–line approximators for nonlinear system identification: a unified approach,” in *Control and Dynamic Systems: Neural Network Systems Techniques and Applications*, X. Leondes, Ed. New York: Academic, 1998, vol. 7, pp. 191–230.
69. M. M. Polycarpou and A. J. Helmicki, “Automated fault detection and accommodation: a learning systems approach,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 25, no. 11, pp. 1447–1458, 1995.
70. M. M. Polycarpou and A. Trunov, “Learning approach to nonlinear fault diagnosis: detectability analysis,” *IEEE Transactions on Automatic Control*, vol. 45, no. 4, pp. 806–812, Apr. 2000.

71. I. Punčochář, J. Široky, and M.Šimandl, “Constrained active fault detection and control,” *IEEE Transactions on Automatic Control*, vol. 60, no. 1, pp. 253–258, 2015.
72. D. M. Raimondo, F. Boem, A. Gallo, and T. Parisini, “A decentralized fault-tolerant control scheme based on active fault diagnosis,” in *IEEE 55th Conference on Decision and Control*, 2016, pp. 2164–2169.
73. D. Raimondo, G. Marseglia, R. Braatz, and J. Scott, “Fault-tolerant model predictive control with active fault isolation,” in *Conference on Control and Fault-Tolerant Systems (SysTol)*. IEEE, 2013, pp. 444–449.
74. R. R. Rajkumar, I. Lee, L. Sha, and J. Stankovic, “Cyber-physical systems: the next computing revolution,” in *Proceedings of the 47th Design Automation Conference*, ser. DAC ’10. New York, NY, USA: ACM, 2010, pp. 731–736.
75. V. Reppa, P. Papadopoulos, M. M. Polycarpou, and C. G. Panayiotou, “A distributed architecture for hvac sensor fault detection and isolation,” *IEEE Transactions on Control Systems Technology*, vol. 23, no. 4, pp. 1323–1337, 2015.
76. V. Reppa, M. M. Polycarpou, and C. G. Panayiotou, “Decentralized isolation of multiple sensor faults in large-scale interconnected nonlinear systems,” *IEEE Transactions on Automatic Control*, vol. 60, no. 6, pp. 1582–1596, 2015.
77. ——, “Distributed sensor fault diagnosis for a network of interconnected cyberphysical systems,” *IEEE Transactions on Control of Network Systems*, vol. 2, no. 1, pp. 11–23, 2015.
78. S. Riverso, F. Boem, G. Ferrari-Trecate, and T. Parisini, “Plug-and-play fault detection and control-reconfiguration for a class of nonlinear large-scale constrained systems,” *IEEE Transactions on Automatic Control*, vol. 61, no. 12, pp. 3963–3978, 2016.
79. T. Samad and T. Parisini, “Systems of Systems,” in *The Impact of Control Technology*, T. Samad and A. M. Annaswamy, Eds. IEEE Control Systems Society, 2011, pp. 175–183. [Online]. Available: ieeecss.org/general/impact-control-technology
80. ——, “Systems of systems,” *The Impact of Control Technology (T.Samad and A.Annaswamy, eds.)*, 2011. [Online]. Available: www.ieeecss.org
81. H. Sandberg, S. Amin, and K. H. Johansson, “Cyberphysical security in networked control systems: An introduction to the issue,” *IEEE Control Systems*, vol. 35, no. 1, pp. 20–23, Feb 2015.
82. J. Scott, R. Findeisen, R. Braatz, and D. Raimondo, “Input design for guaranteed fault diagnosis using zonotopes,” *Automatica*, vol. 50, no. 6, pp. 1580–1589, 2014.
83. L. Sha, S. Gopalakrishnan, X. Liu, and Q. Wang, “Cyber-physical systems: A new frontier,” in *Machine Learning in Cyber Trust*. Springer US, 2009, pp. 3–13.
84. I. Shames, A. M. Teixeira, H. Sandberg, and K. H. Johansson, “Distributed fault detection for interconnected second-order systems,” *Automatica*, vol. 47, no. 12, pp. 2757–2764, 2011.
85. F. Shi and R. Patton, “Fault estimation and active fault tolerant control for linear parameter varying descriptor systems,” *International Journal of Robust and Nonlinear Control*, vol. 25, no. 5, pp. 689–706, 2015.
86. M. Simandl and I. Puncochar, “Active fault detection and control: Unified formulation and optimal design,” *Automatica*, vol. 45, no. 9, pp. 2052–2059, 2009.
87. J. Škach, I. Punčochář, and F. L. Lewis, “Optimal active fault diagnosis by temporal-difference learning,” in *Decision and Control (CDC), 2016 IEEE 55th Conference on*. IEEE, 2016, pp. 2146–2151.
88. R. S. Smith, “Covert misappropriation of networked control systems: Presenting a feedback structure,” *IEEE Control Systems*, vol. 35, no. 1, pp. 82–92, 2015.
89. S. Stankovic, N. Ilic, Z. Djurovic, M. Stankovic, and K. Johansson, “Consensus based overlapping decentralized fault detection and isolation,” in *Conference on Control and Fault-Tolerant Systems (SysTol’10)*, 2010, pp. 570–575.
90. M. Staroswiecki and A. M. Amani, “Fault-tolerant control of distributed systems by information pattern reconfiguration,” *International Journal of Adaptive Control and Signal Processing*, 2014.
91. M. Staroswiecki and A. Amani, “Fault-tolerant control of distributed systems by information pattern reconfiguration,” *International Journal of Adaptive Control and Signal Processing*, vol. 29, no. 6, pp. 671–684, 2015.

92. S. Tabatabaeipour, “Active fault detection and isolation of discrete-time linear time-varying systems: a set-membership approach,” *International Journal of Systems Science*, vol. 46, no. 11, pp. 1917–1933, 2015.
93. P. Tabuada, “Cyber physical systems: Position paper,” in *NSF Workshop on Cyber-Physical Systems*, 2006.
94. P. L. Tang and C. de Silva, “Compensation for transmission delays in an ethernet-based control network using variable-horizon predictive control,” *IEEE Transactions on Control Systems Technology*, vol. 14, no. 4, pp. 707 – 718, 2006.
95. A. Teixeira, H. Sandberg, and K. Johansson, “Networked control systems under cyber attacks with applications to power networks,” in *American Control Conference (ACC), 2010*, 30 2010-july 2 2010, pp. 3690 –3696.
96. A. Teixeira, I. Shames, H. Sandberg, and K. H. Johansson, “Distributed fault detection and isolation resilient to network model uncertainties,” *IEEE Transactions on Cybernetics*, vol. 44, no. 11, pp. 2024–2037, Nov 2014.
97. R. Tempo, G. Calafiore, and F. Dabbene, *Randomized algorithms for analysis and control of uncertain systems: with applications*. Springer Science & Business Media, 2012.
98. A. Vemuri and M. M. Polycarpou, “On-line approximation methods for robust fault detection,” *Proc. 13th IFAC World Congress*, vol. K, pp. 319–324, 1996.
99. V. Venkatasubramanian, R. Rengaswamy, S. Kavuri, and K. Yin, “A review of process fault detection and diagnosis:: Part III: Process history based methods,” *Computers & Chemical Engineering*, vol. 27, no. 3, pp. 327–346, Mar. 2003.
100. V. Venkatasubramanian, R. Rengaswamy, K. Yin, and S. Kavuri, “A review of process fault detection and diagnosis Part I: Quantitative model-based methods,” *Computers & Chemical Engineering*, vol. 27, no. 3, pp. 293–311, Mar. 2003.
101. V. Venkatasubramanian, R. Rengaswamy, and S. Kavuri, “A review of process fault detection and diagnosis:: Part II: Qualitative models and search strategies,” *Computers & Chemical Engineering*, vol. 27, no. 3, pp. 313–326, 2003.
102. L. Wei, W. Gui, Y. Xie, and S. X. Ding, “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*, 2009, pp. 816–821.
103. W. Wolf, “Cyber-physical systems,” *Computer*, vol. 42, no. 3, pp. 88 –89, march 2009.
104. F. Xu, S. Olaru, V. Puig, C. Ocampo-Martinez, and S.-I. Niculescu, “Sensor-fault tolerance using robust mpc with set-based state estimation and active fault isolation,” *International Journal of Robust and Nonlinear Control*, 2016.
105. F. Xu, V. Puig, C. Ocampo-Martinez, and X. Wang, “Set-valued observer-based active fault-tolerant model predictive control,” *Optimal Control Applications and Methods*, 2016.
106. S. Zanero, “Cyber-physical systems,” *Computer*, vol. 50, no. 4, pp. 14–16, 2017.
107. X. Zhang, M. M. Polycarpou, and T. Parisini, “A robust detection and isolation scheme for abrupt and incipient faults in nonlinear systems,” vol. 47, no. 4, pp. 576–593, 2002.
108. ——, “Decentralized fault detection for a class of large-scale nonlinear uncertain systems,” in *48h IEEE Conference on Decision and Control and 28th Chinese Control Conference*, 2009, pp. 6988–6993.
109. K. Zhou and J. C. Doyle, *Essentials of robust control*. Prentice hall Upper Saddle River, NJ, 1998, vol. 104.
110. Y. Zhou, F. Boem, C. Fischione, and T. Parisini, “Distributed fault detection with sensor networks using pareto-optimal dynamic estimation method,” in *2016 European Control Conference (ECC)*, 2016, pp. 728–733.