Distributed Sensor Fault Diagnosis for a Network of Interconnected Cyber-Physical Systems

Vasso Reppa, *Member, IEEE,* Marios M. Polycarpou, *Fellow, IEEE,* and Christos G. Panayiotou, *Senior Member, IEEE*

***Abstract*—This paper proposes a distributed methodology for detecting and isolating multiple sensor faults in interconnected cyber-physical systems. The distributed sensor fault detection and isolation process is conducted in the cyber superstratum, in two levels. The first-level diagnosis is based on the design of monitoring agents, where every agent is dedicated to a corresponding interconnected subsystem. The monitoring agent is designed to isolate multiple sensor faults occurring in the sensor set of the physical part, while it is allowed to exchange information with its neighboring monitoring agents. The second-level diagnosis is realized by applying a global decision logic designed to isolate multiple sensor faults that may propagate in the cyber superstratum through the exchange of information between monitoring agents. The decision making process, executed in both levels of diagnosis, relies on a multiple sensor fault combinatorial logic and diagnostic reasoning. The performance of the proposed methodology is analyzed with respect to the sensor fault propagation effects and the distributed sensor fault detectability.**

I. Introduction

Recent advances in information and communication technologies, embedded systems and sensor networks have generated significant research activity in the development of the so-called cyber-physical systems (CPS). According to [1], CPS consist of (i) physical, biological or engineered systems that are usually large-scale and complex, and (ii) a cyber core, comprised of communication networks and computational availability that monitors, coordinates and controls the physical part. The focus of CPS is to improve the collaborative link between physical and computational (cyber) elements for increased adaptability, efficiency and autonomy. The key motivation for migrating from "conventional" systems to CPS is the need for enhancing the "intelligence" of the physical systems used in many application domains in order to be able to plan and modify their actions based on self-awareness and the evolving environment, and for handling a huge amount of data of different time and space characteristics.

Among the key challenges in designing CPS are safety, reliability and fault tolerance. For meeting these challenges, the cyber core should be empowered with supervision capabilities for diagnosing faults in the physical part and compensating their effects by taking appropriate remedial actions [2], [3]. Various methodologies have been developed for tackling the problem of fault diagnosis in the framework of actuator/sensor faults and especially process faults. Recently, the detection and isolation of sensor faults have become of paramount importance, mainly as a result of the large number of sensors and sensor networks, used for (i) monitoring and controlling large-scale CPS; (ii) providing rich and redundant information for executing safety-critical tasks; and (iii) offering information to the citizens and governmental agencies for resolving problems promptly in emergency situations.

Emerging applications of CPS with multiple sensors can be found in intelligent transportation, smart buildings, smart grids, mobile robotics and many more. For instance, in intelligent transportation, vehicles may be equipped with odometers, lasers, frontal camera video-sensors, GPS, speed or object tracking sensors, in order to be able to acquire and broadcast information aiming at performing tasks such as cooperative or fully autonomous driving, avoiding lane departure and collision, etc. In smart buildings, multiple sensors are installed in different zones (e.g. temperature sensor, humidity sensor, CO_2 sensor, contaminant concentration sensor, infrared occupancy sensor), as well as in the electromechanical part of heating, ventilation and air-conditioning systems for measuring supply/return/mixed air temperature, supply/return air differential pressure, return air humidity, etc. Such sensing information may be used for reducing the energy consumption of a building and maintaining the desired living conditions, while executing evacuation plans in safety-critical situations (e.g. fire). Any undetected sensor faults can have severe consequences, possibly leading to system instability, loss of information fidelity, incorrect decisions and disorientation of remedial actions.

Due to the large-scale and complex nature of the physical systems, it is convenient to model CPS as a set of interconnected subsystems of lower dimension. In this context, decentralized and distributed approaches are commonly adopted for online fault detection and isolation (FDI). The common characteristic of model-based decentralized and distributed FDI schemes is the development of local monitoring units that perform diagnosis based on local models describing the interconnected subsystems. The classification of these schemes relies on the type of system interconnections, the cyber levels of diagnosis, the task of the local diagnosers, as well as the type of communication and information exchanged between the local and high-level diagnosers. The design of model-based

V. Reppa is with the Automatic Control Department, Supélec, Gif-Sur-Yvette, 91400, France (e-mail: vasiliki.reppa@supelec.fr, tel: +33169851384). M. M. Polycarpou and C. G. Panayiotou are with the KIOS Research Center for Intelligent Systems and Networks, Department of Electrical and Computer Engineering, University of Cyprus, Nicosia 1678, Cyprus (e-mail: {mpolycar,christosp}@ucy.ac.cy, tel: +3522893451). This work is funded by the European Research Council Advanced Grant FAULT-ADAPTIVE (ERC-AdG-291508) and People Programme (Marie Curie Actions) of the European Union's 7th Framework Programme (FP7/2007-2013) under REA grant agreement n° [626891].

techniques for interconnected systems requires some prior knowledge related to the subsystems and the interconnections.

In [4]–[10], decentralized and distributed FDI methods are developed for physically interconnected subsystems. Distributed architectures have also been designed for systems with interconnections in the control law [11], interconnected inputs [12] or sensing interconnections (i.e. relative output measurements) [13]. For enhancing fault isolation, multiple levels of diagnosis have been designed. In particular, while the single level diagnosis is realized by the local diagnosers [4]–[6], [11]–[13], additional FDI units are developed, aggregating and processing the outputs of the local diagnosers [7]–[10]. The decentralized or distributed nature of the FDI process is related either to the task executed by the local diagnosers or the communication between the local diagnosers. In decentralized schemes, a local diagnoser is commonly designed to detect and isolate faults only in its underlying system [4], [6], while it may not exchange any information with other local diagnosers [9], [10]. On the contrary, in distributed schemes, there is communication between the local diagnosers and every local diagnoser can detect and isolate faults in neighboring systems [5], [7], [8], [11]–[13]. The design of distributed FDI architectures may also differ in the type of exchanged information. Specifically, the local diagnosers may exchange estimations [4], [5], [12], or measurements [7], [8], [11] of the interconnected states, or fault signatures [12]. In multi-level FDI schemes, the communication between levels is commonly sporadic and event-driven, while the information transmitted to higher levels can be the decisions of the local diagnosers [7], [8], the time instances of fault detection of the local diagnosers [9] or the calculated analytical redundancy relations [10].

Among distributed and decentralized FDI schemes for physically interconnected systems, there are very few results in sensor fault diagnosis. In [14], [15], a distributed architecture has been designed for isolating a single sensor fault that may occur in one of the nonlinear subsystems. In [16], [17], local monitoring agents, which do not exchange any information, are used for isolating multiple sensor faults that may affect more than one interconnected, nonlinear subsystem.

The goal and the main contribution of this work is the design and analysis of a fault diagnosis methodology with emphasis on the distributed isolation of *multiple sensor faults* that may affect the physical part of *multiple* interconnected cyber-physical systems, which may exchange sensor information related to the physical interconnections. The backbone of the proposed distributed scheme is the use of a bank of agents, which are implemented in the cyber core of the interconnected CPS and monitor the sensor sets of the CPS. A global decision logic is designed for the isolation of sensor faults that are propagated between the interconnected CPS through the exchange of sensor information, which is necessary for enhancing the distributed sensor fault detectability. This exchange of information is crucial and has significant practical implications, since it provides the necessary redundancy to isolate multiple faults in large-scale dynamical systems. The intuitive rationale behind the global decision logic relies on the diagnosis capabilities of a monitoring agent, which is specifically designed using a bank of observer-based modules that are robust against modeling uncertainties and structurally sensitive to propagated sensor faults and faults that occur in smaller local sensor sets of the underlying CPS, while being affected by local sensor faults in a different way than being affected by propagated sensor faults.

The local sensor sets result from the decomposition of the sensor set of the corresponding CPS, necessary for resolving efficiently the problem of isolating multiple sensor faults in possibly large-scale and nonlinear CPS, and distinguishing propagated sensor faults. The decision logic of a monitoring module relies on analytical redundancy relations of residuals and adaptive thresholds, derived using a nonlinear Lipschitz observer. The design of the nonlinear observer is realized by taking into account certain conditions that ensure the stability of the nonlinear estimation error dynamics, using the measurements of the underlying local sensor set and the sensor information transmitted from neighboring CPS. These conditions are used to analyze quantitatively the distributed sensor fault effects.

The isolation decision logic applied locally, by combining the decisions of the monitoring modules, and globally, by combining the decisions of the monitoring agents of the CPS, relies on diagnostic reasoning using sensor fault signature matrices. The proposed distributed diagnostic reasoning and fault signature matrices are formulated taking into account the robustness and structured fault sensitivity properties, as well as the quantitative analysis of the local and propagated sensor fault effects, allowing the isolation of multiple sensor faults in a CPS and multiple propagated sensor faults that impact the network of the interconnected CPS. This work is based on some preliminary results on distributed sensor fault diagnosis in [18], while offering a general design and analysis framework for diagnosing multiple sensor faults that may affect the physical layer of operation of a network of CPS and propagate in the cyber superstratum.

The paper is organized as follows. The problem formulation is described in Section II, while the overall architecture of the distributed sensor FDI method for a network of CPS is presented in Section III. The distributed sensor fault detection and isolation procedures are described in Section IV and V, respectively. The performance of the proposed methodology is analyzed in Section VI. A simple two-zone Heating Ventilation and Air-Conditioning (HVAC) system is provided in Section VII for illustrating the application of the proposed method, followed by some concluding remarks in Section VIII.

II. Problem Formulation

Consider a network of N interconnected CPS. The I-th CPS, $I \in \{1, \ldots, N\}$, is described by the pair $\left(\mathcal{P}^{(I)}, \mathcal{C}^{(I)}\right)$, where $\mathcal{P}^{(I)}$ corresponds to the physical part, while $\mathcal{C}^{(I)}$ denotes the cyber part. The physical part $\mathcal{P}^{(I)}$ is modeled by a nonlinear dynamical subsystem, denoted by $\Sigma^{(I)}$, and a set of sensors, denoted by $\mathcal{S}^{(I)}$; i.e.,

$$\begin{aligned}\Sigma^{(I)}: \quad \dot{x}^{(I)}(t) &= A^{(I)}x^{(I)}(t) + \gamma^{(I)}(x^{(I)}(t), u^{(I)}(t)) \\ &\quad + h^{(I)}(x^{(I)}(t), u^{(I)}(t), C_z^{(I)} z^{(I)}(t)) \\ &\quad + \eta^{(I)}(x^{(I)}(t), u^{(I)}(t), C_z^{(I)} z^{(I)}(t), t), \quad (1)\end{aligned}$$

where $x^{(I)} \in \mathbb{R}^{n_I}$, $u^{(I)} \in \mathbb{R}^{\ell_I}$ are the state and control input vector of $\Sigma^{(I)}$, respectively, while $z^{(I)} \in \mathbb{R}^{q_I}$ is the interconnection state vector, containing the states of the neighboring (interconnected) subsystems of $\Sigma^{(I)}$, while $C_z^{(I)} z^{(I)}$ denotes a linear combination of interconnection states with $C_z^{(I)} \in \mathbb{R}^{p_I \times q_I}$. The constant matrix $A^{(I)} \in \mathbb{R}^{n_I \times n_I}$ is the linearized part of the state equation and $\gamma^{(I)} : \mathbb{R}^{n_I} \times \mathbb{R}^{\ell_I} \mapsto \mathbb{R}^{n_I}$ represents the known nonlinear dynamics. The term $A^{(I)}x^{(I)} + \gamma^{(I)}(x^{(I)}, u^{(I)})$ represents the known local dynamics, while $h^{(I)} : \mathbb{R}^{n_I} \times \mathbb{R}^{\ell_I} \times \mathbb{R}^{p_I} \mapsto \mathbb{R}^{n_I}$ represents the known interconnection dynamics. The last term $\eta^{(I)} : \mathbb{R}^{n_I} \times \mathbb{R}^{\ell_I} \times \mathbb{R}^{p_I} \times \mathbb{R} \mapsto \mathbb{R}^{n_I}$ denotes the modeling uncertainty of $\Sigma^{(I)}$, representing various sources of uncertainty such as linearization error, uncertainty in the model's parameters, or system disturbances etc. The input vector $u^{(I)}$ is generated by a control agent, denoted by $\mathcal{K}^{(I)}$, which is implemented in the cyber part $\mathcal{C}^{(I)}$ of the I-th CPS, based on some desired reference signals $r^{(I)}(t)$.

The sensor set of $\mathcal{P}^{(I)}$ is characterized by

$$\mathcal{S}^{(I)}: \qquad y^{(I)}(t) = C^{(I)}x^{(I)}(t) + d^{(I)}(t) + f^{(I)}(t), \quad (2)$$

where $y^{(I)} \in \mathbb{R}^{m_I}$ is the output vector, $d^{(I)} \in \mathbb{R}^{m_I}$ denotes the noise vector corrupting the measurements of sensors in $\mathcal{S}^{(I)}$ and $f^{(I)} \in \mathbb{R}^{m_I}$ represents the possible sensor fault vector. The j-th sensor $\mathcal{S}^{(I)}\{j\}$, $j \in \{1, \ldots, m_I\}$ is described by

$$\mathcal{S}^{(I)}\{j\}: \qquad y_j^{(I)}(t) = C_j^{(I)}x^{(I)}(t) + d_j^{(I)}(t) + f_j^{(I)}(t), \quad (3)$$

where $f_j^{(I)}$ represents the change in the output $y_j^{(I)}$ due to a single fault in the j-th sensor of the set $\mathcal{S}^{(I)}$, modeled as

$$f_j^{(I)}(t) = \beta_j^{(I)}(t - T_{f_j}^{(I)})\phi_j^{(I)}(t - T_{f_j}^{(I)}), \quad (4)$$

where $\beta_j^{(I)}(t)$ is the time profile and $\phi_j^{(I)}(t)$ is the (unknown) sensor fault function that occurs at the (unknown) time instant $T_{f_j}^{(I)}$. The time profile of the fault is modeled as $\beta_j^{(I)}(t) = 0$ if $t < 0$ and $\beta_j^{(I)}(t) = 1 - e^{-\kappa_j^{(I)} t}$ if $t \geq 0$, where $\kappa_j^{(I)} > 0$ denotes the (unknown) evolution rate of the fault. If the occurrence of a sensor fault is abrupt, $\kappa_j^{(I)} \to \infty$. Multiple sensor faults may occur simultaneously or sequentially; e.g., $T_{f_1}^{(I)} \leq T_{f_2}^{(I)} \leq \ldots \leq T_{f_{m_I}}^{(I)}$.

In the presence of a single or multiple sensor faults, it is important that the monitoring system is able to diagnose the faults as quickly and accurately as possible before they lead to catastrophic failures or propagate to other CPS through the distributed control scheme. The objective of this work is to design and analyze a methodology for detecting and isolating multiple sensor faults that may occur in one or more CPS. The following assumptions are made throughout the paper:

Assumption 1: For each subsystem $I \in \{1, \ldots, N\}$, the state vector $x^{(I)}(t)$ and input vector $u^{(I)}(t)$ generated by a feedback controller $\mathcal{K}^{(I)}$, remain uniformly bounded before and after the occurrence of multiple sensor faults; i.e., there exist compact regions of stability $\mathcal{U}_I \subset \mathbb{R}^{\ell_I}$, $\mathcal{X}_I \subset \mathbb{R}^{n_I}$ such that $\big(x^{(I)}(t), u^{(I)}(t)\big) \in \mathcal{X}_I \times \mathcal{U}_I$, for all $t \geq 0$.

Assumption 2: The nonlinear vector field $\gamma^{(I)}$ is locally Lipschitz in $x^{(I)} \in \mathcal{X}_I$, for all $u^{(I)} \in \mathcal{U}_I$ and $t \geq 0$, while $h^{(I)}$ is locally Lipschitz in $x^{(I)} \in \mathcal{X}_I$ and $z^{(I)} \in \mathcal{Z}_I$, for all $u^{(I)} \in \mathcal{U}_I$ and $t \geq 0$. The vector space $\mathcal{Z}_I \subset \mathbb{R}^{q_I}$ denotes a compact region of stability within which $z^{(I)}$ resides.

Assumption 3: The unknown modeling uncertainty $\eta^{(I)}$, is bounded by a known functional bound $\bar{\eta}^{(I)}$ for all $x^{(I)} \in \mathcal{X}_I$, $u^{(I)} \in \mathcal{U}_I$, $z^{(I)} \in \mathcal{Z}_I$ and $t \geq 0$; i.e.

$$|\eta^{(I)}(x^{(I)}, u^{(I)}, C_z^{(I)}z^{(I)}, t)| \leq \bar{\eta}^{(I)}(x^{(I)}, u^{(I)}, C_z^{(I)}z^{(I)}, t), \quad (5)$$

whereas the functional bound $\bar{\eta}^{(I)}$ is locally Lipschitz in $x^{(I)} \in \mathcal{X}_I$ and $z^{(I)} \in \mathcal{Z}_I$, for all $u^{(I)} \in \mathcal{U}_I$ and $t \geq 0$.

Assumption 4: The noise affecting the sensor $\mathcal{S}^{(I)}\{j\}$ is unknown but uniformly bounded; i.e., $|d_j^{(I)}(t)| \leq \bar{d}_j^{(I)}$, $j \in \{1, \ldots, m_I\}$, where $\bar{d}_j^{(I)}$ is a known constant bound.

Assumption 1 is a well-posedness condition, requiring that the feedback controller can retain the boundedness of the state variables in the presence of sensor faults. This assumption is necessary due to the fact that, in this work, we do not address the fault accommodation problem, but only fault detection and isolation issues. Assumption 2 characterizes the class of nonlinear interconnected systems under consideration. Many nonlinearities in practical systems can be considered as locally Lipschitz [4], [19], [20]. Assumption 3 provides a bound commonly used for distinguishing between modeling uncertainties and faults. This bound can be obtained either analytically, by explicitly determining the sources of uncertainty and their corresponding bounds, or using off-line identification techniques. Assumption 4 describes a practical representation of the available knowledge for the sensor noise that is typically provided in a given range of operation by sensor manufacturers or introduced when a noise-free analog signal is converted into a digital one with a finite number of digits.

III. Sensor Fault Diagnosis Architecture

The cyber part $\mathcal{C}^{(I)}$ of the interconnected CPS consists of the monitoring agent, denoted by $\mathcal{M}^{(I)}$, associated with each interconnected subsystem and the control agent $\mathcal{K}^{(I)}$. The objective of this paper is the design of the monitoring agent $\mathcal{M}^{(I)}$ for multiple sensor faults, with emphasis on the communication between neighboring monitoring agents and the decision logic for sensor fault isolation. The overall architecture of the proposed distributed sensor fault detection and isolation (SFDI) method is illustrated in Fig. 1. The distributed SFDI process is realized in the cyber superstratum, in two levels. The first-level diagnosis $\mathcal{M}^{(I)}$, shown in Fig. 2, is designed to detect and isolate multiple sensor faults that may affect directly some sensors in $\mathcal{S}^{(I)}$ associated with the underlying system $\Sigma^{(I)}$. The second-level diagnosis, denoted by $\mathcal{G}$, utilizes information (depicted as dashed arrows in Fig. 1) from the monitoring agents $\mathcal{M}^{(I)}$ for isolating sensor faults propagated in the cyber superstratum due to the communication of the monitoring agents $\mathcal{M}^{(I)}$.

The exchanged information corresponds to the form of the physical interconnections, i.e. unidirectional or bidirectional interactions between interconnected subsystems (white, solid arrows in Fig. 1). In particular, neighboring monitoring agents can exchange data provided by the sensors that measure the interconnection states $z^{(I)}$. The sensor information transmitted

to the agent $\mathcal{M}^{(I)}$, denoted by $\mathcal{S}_z^{(I)}$, is characterized by the output vector $y_z^{(I)} \in \mathbb{R}^{p_I}$; i.e.

$$\mathcal{S}_z^{(I)}: \; y_z^{(I)}(t) = C_z^{(I)} z^{(I)}(t) + d_z^{(I)}(t) + f_z^{(I)}(t), \quad (6)$$

where $d_z^{(I)}$, $f_z^{(I)} \in \mathbb{R}^{p_I}$ are the noise and sensor fault vector, respectively.

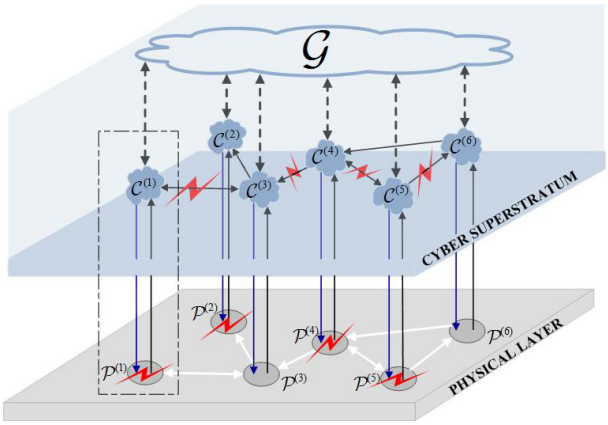


Fig. 1. Network of interconnected cyber-physical systems.

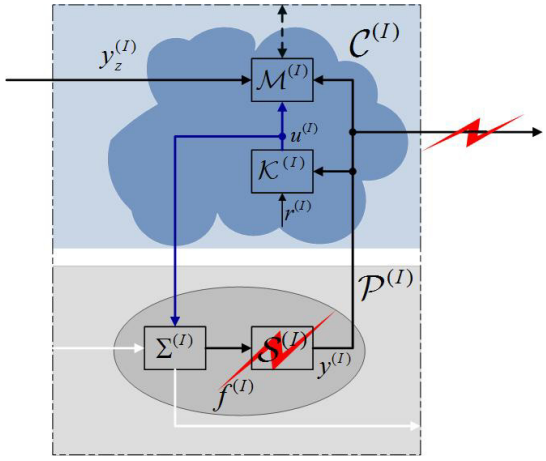


Fig. 2. Description of the I-th CPS affected by sensor faults.

A fault in sensor $\mathcal{S}^{(I)}$ can be propagated to neighboring cyber modules due to the information exchange of $\mathcal{M}^{(I)}$ with neighboring monitoring agents, as depicted in Fig. 2. Allowing the exchange of sensor information between monitoring agents can enhance sensor fault detectability, compared to a decentralized architecture with no communication between the monitoring agents, where usually the effects of interconnections are treated as bounded disturbances [16], [17]. On the other hand, the exchange of information may cause fault propagation, which complicates the isolation of faulty sensors. This issue is addressed in Section V with the design of a global decision logic. It is important to note that the propagation of sensor faults may occur not only in the monitoring agents but also in the control agents in the case of a distributed control architecture, however the design of the control and fault accommodation scheme is beyond the scope of this paper.

Typically, in large-scale applications, the sensor set $\mathcal{S}^{(I)}$ may consist of a large number of sensors, thus making the detection and especially the isolation of multiple sensor faults very difficult and sometimes unfeasible with a single module. The design of the agent $\mathcal{M}^{(I)}$ relies on the decomposition of the monitoring of the sensor set $\mathcal{S}^{(I)}$ into smaller local sensor sets that may be distinct or overlap [17]. The q-th local sensor set, denoted by $\mathcal{S}^{(I,q)}$, consists of $m_{I,q}$ sensors of $\mathcal{S}^{(I)}$ and is characterized by the output vector $y^{(I,q)} \in \mathbb{R}^{m_{I,q}}$; i.e.,

$$\mathcal{S}^{(I,q)}: \quad y^{(I,q)}(t) = C^{(I,q)} x^{(I)}(t) + d^{(I,q)}(t) + f^{(I,q)}(t), \quad (7)$$

where $C^{(I,q)}$ is made up of $m_{I,q}$ rows of $C^{(I)}$ and $y^{(I,q)}$ represents a column vector made up of $m_{I,q}$ elements of $y^{(I)}$ (correspondingly for $d^{(I,q)}$ and $f^{(I,q)}$). A decomposition procedure that can ensure the design of stable observers of the monitoring modules (presented in Section IV-A) and enhance the multiple sensor fault isolability is proposed in [17].

For each local sensor set, we design a dedicated module $\mathcal{M}^{(I,q)}$, which uses the measurements of $\mathcal{S}^{(I,q)}$, as well as the sensor information $y_z^{(I)}$ transmitted to the agent $\mathcal{M}^{(I)}$ from its neighboring agents. The primary goal of $\mathcal{M}^{(I,q)}$ is to detect the presence of sensor faults $f^{(I,q)}$ affecting the local sensor set $\mathcal{S}^{(I,q)}$. However, each module $\mathcal{M}^{(I,q)}$ uses the transmitted sensor information $y_z^{(I)}$, which may be faulty, thus affecting the decision of $\mathcal{M}^{(I,q)}$; i.e., the module $\mathcal{M}^{(I,q)}$ cannot distinguish between sensor faults in $\mathcal{P}^{(I)}$ and propagated sensor faults. Due to this fact, as well as due to the possible overlapping between some local sensor sets $\mathcal{S}^{(I,q)}$ (sensors belong to more than one local sensor sets), the decisions of the N_I modules are aggregated and processed combinatorially, applying diagnostic reasoning, in order for the agent $\mathcal{M}^{(I)}$ to isolate multiple sensor faults. Then, the decisions of the monitoring agents are processed by a global decision logic, aiming at isolating propagated sensor faults.

IV. Distributed Sensor Fault Detection

This section deals with the design of the module $\mathcal{M}^{(I,q)}$, $q \in \{1, \ldots, N_I\}$. In the sequel, the dependence of the signals on time (e.g. $x(t)$) will be dropped for notational brevity.

A. Observer-based residual generation

The estimation model of $\mathcal{M}^{(I,q)}$ is formulated by selecting a nonlinear observer $\mathscr{O}^{(I,q)}$, described by

$$\begin{aligned} \mathscr{O}^{(I,q)}: \quad \dot{\hat{x}}^{(I,q)} &= A^{(I)} \hat{x}^{(I,q)} + \gamma^{(I)}(\hat{x}^{(I,q)}, u^{(I)}) \\ &\quad + h^{(I)}(\hat{x}^{(I,q)}, u^{(I)}, y_z^{(I)}) \\ &\quad + L^{(I,q)} \left(y^{(I,q)} - C^{(I,q)} \hat{x}^{(I,q)} \right), \end{aligned} \quad (8)$$

where $\hat{x}^{(I,q)} \in \mathbb{R}^{n_I}$ is the estimation of $x^{(I)}$ (based on the sensor measurements $y^{(I,q)}$ with $\hat{x}^{(I,q)}(0) = 0$), $L^{(I,q)} \in \mathbb{R}^{n_I \times m_{I,q}}$ is the observer gain matrix and $y_z^{(I)}$ is the transmitted sensor information, defined in (6). It is noted that the observer used in this work is based on the structure of observers for Lipschitz nonlinear systems (see [20], [21] and references therein), which is modified appropriately for non-

linear interconnected subsystems. The j-th residual, denoted by $\varepsilon_{y_j}^{(I,q)}$, is defined as

$$\varepsilon_{y_j}^{(I,q)} = y_j^{(I)} - C_j^{(I)}\hat{x}^{(I,q)}, \quad j \in \mathcal{J}^{(I,q)}, \tag{9}$$

where $\mathcal{J}^{(I,q)}$ is an index set, defined as $\mathcal{J}^{(I,q)} = \{j : \mathcal{S}^{(I)}\{j\} \in \mathcal{S}^{(I,q)}\}$.

As long as the local sensor set $\mathcal{S}^{(I,q)}$ and the transmitted information $y_z^{(I)}$ are not affected by sensor faults, the estimation model of $\mathcal{M}^{(I,q)}$ under healthy conditions, denoted by $\hat{x}_H^{(I,q)}$, satisfies

$$\begin{aligned}\dot{\hat{x}}_H^{(I,q)} =& A_L^{(I,q)}\hat{x}_H^{(I,q)} + \gamma^{(I)}(\hat{x}_H^{(I,q)}, u^{(I)}) \\ &+ h^{(I)}(\hat{x}_H^{(I,q)}, u^{(I)}, y_{z_H}^{(I)}) + L^{(I,q)}y_H^{(I,q)},\end{aligned} \tag{10}$$

where $A_L^{(I,q)} = A^{(I)} - L^{(I,q)}C^{(I,q)}$ and $y_H^{(I,q)}$ and $y_{z_H}^{(I)}$ are respectively defined through (7) and (6) with $f^{(I,q)} = 0$ and $f_z^{(I,q)} = 0$.

Let us define $\varepsilon_{x_H}^{(I,q)} \triangleq x^{(I)} - \hat{x}_H^{(I,q)}$ as the state estimation error under healthy conditions; taking into account (1) and (10), we obtain:

$$\begin{aligned}\dot{\varepsilon}_{x_H}^{(I,q)} =& A_L^{(I,q)}\varepsilon_{x_H}^{(I,q)} + \tilde{\gamma}_H^{(I,q)} + \tilde{h}_H^{(I,q)} - L^{(I,q)}d^{(I,q)} \\ &+ \eta^{(I)}(x^{(I)}, u^{(I)}, C_z^{(I)}z^{(I)}, t),\end{aligned} \tag{11}$$

$$\tilde{\gamma}_H^{(I,q)} = \gamma^{(I)}(x^{(I)}, u^{(I)}) - \gamma^{(I)}(\hat{x}_H^{(I,q)}, u^{(I)}), \tag{12}$$

$$\tilde{h}_H^{(I,q)} = h^{(I)}(x^{(I)}, u^{(I)}, C_z^{(I)}z^{(I)}) - h^{(I)}(\hat{x}_H^{(I,q)}, u^{(I)}, y_{z_H}^{(I)}) \tag{13}$$

The stability of the estimation error dynamics under healthy conditions is analyzed in the following lemma.

Lemma 4.1: Suppose that the observer gain $L^{(I,q)}$ is chosen such that: (a) the matrix $A_L^{(I,q)} = A^{(I)} - L^{(I,q)}C^{(I,q)}$ is stable, and (b) there exist positive constants $\rho^{(I,q)}$, $\xi^{(I,q)}$ such that $|e^{A_L^{(I,q)}t}| \leq \rho^{(I,q)}e^{-\xi^{(I,q)}t}$ and $\xi^{(I,q)} > \Lambda_I\rho^{(I,q)}$, where $\Lambda_I = \lambda_{\gamma_I} + \lambda_{h_I} + \lambda_{\eta_I}$ (the parameters λ_{γ_I}, λ_{h_I}, λ_{η_I} denote the Lipschitz constants of $\gamma^{(I)}$, $h^{(I)}$ and $\bar{\eta}^{(I)}$, respectively); then the state estimation error under healthy conditions, $\varepsilon_{x_H}^{(I,q)}(t)$, is uniformly bounded and satisfies

$$|\varepsilon_{x_H}^{(I,q)}(t)| \leq E_H^{(I,q)}(t) + \rho^{(I,q)}\Lambda_I \int_0^t E_H^{(I,q)}(\tau)e^{-\nu^{(I,q)}(t-\tau)}d\tau \tag{14}$$

$$\begin{aligned}E_H^{(I,q)}(t) =& \Phi(t)\bar{x}^{(I)} + \frac{\rho_d^{(I,q)}\bar{d}^{(I,q)}}{\xi_d^{(I,q)}}\left(1 - e^{-\xi_d^{(I,q)}t}\right) \\ &+ \int_0^t \Phi(t-\tau)\bar{\eta}^{(I)}(\hat{x}_H^{(I,q)}(\tau), u^{(I)}(\tau), y_{z_H}^{(I)}(\tau), \tau)d\tau \\ &+ \frac{(\lambda_{h_I} + \lambda_{\eta_I})\bar{d}_z^{(I)}}{\xi^{(I,q)}}\left(\rho^{(I,q)} - \Phi(t)\right).\end{aligned} \tag{15}$$

with $\Phi(t) = \rho^{(I,q)}e^{-\xi^{(I,q)}t}$ and $\nu^{(I,q)} = \xi^{(I,q)} - \rho^{(I,q)}\Lambda_I$.

Proof: Solving (11) and taking into account conditions (a) and (b) of Lemma 4.1, the bound on $\left|\varepsilon_{x_H}^{(I,q)}(t)\right|$ satisfies:

$$\begin{aligned}\left|\varepsilon_{x_H}^{(I,q)}(t)\right| \leq& \Phi(t)\bar{x}^{(I)} + \int_0^t \rho_d^{(I,q)}e^{-\xi_d^{(I,q)}(t-\tau)}\bar{d}^{(I,q)}d\tau \\ &+ \int_0^t \Phi(t-\tau)\left(\left|\tilde{\gamma}_H^{(I,q)}(\tau)\right| + \left|\tilde{h}_H^{(I,q)}(\tau)\right|\right. \\ &\left. + \left|\eta^{(I)}(x^{(I)}(\tau), u^{(I)}(\tau), C_z^{(I)}z^{(I)}(\tau), \tau)\right|\right)d\tau,\end{aligned} \tag{16}$$

where $\rho_d^{(I,q)}$, $\xi_d^{(I,q)}$ are positive constants chosen such that $\left|e^{A_L^{(I,q)}t}L^{(I,q)}\right| \leq \rho_d^{(I,q)}e^{-\xi_d^{(I,q)}t}$, $\bar{x}^{(I)}$ is a bound on $\left|x^{(I)}(t)\right|$ such that $\left|x^{(I)}(t)\right| \leq \bar{x}^{(I)}$, for all $x^{(I)} \in \mathcal{X}^{(I)}$ and $t \geq 0$, $\bar{d}^{(I,q)}$ is a bound for $d^{(I,q)}(t)$ (defined in (7)) such that $\left|d^{(I,q)}(t)\right| \leq \bar{d}^{(I,q)}$, and $\tilde{\gamma}_H^{(I,q)}$, $\tilde{h}_H^{(I,q)}$ are defined in (12) and (13), respectively. Given Assumptions 1-4, we have

$$\left|\tilde{\gamma}_H^{(I,q)}\right| \leq \lambda_{\gamma_I}\left|\varepsilon_{x_H}^{(I,q)}\right|, \tag{17}$$

$$\left|\tilde{h}_H^{(I,q)}\right| \leq \lambda_{h_I}\left|\begin{bmatrix}\varepsilon_{x_H}^{(I,q)} \\ C_z^{(I)}z^{(I)} - y_{z_H}^{(I)}\end{bmatrix}\right| \leq \lambda_{h_I}\left(\left|\varepsilon_{x_H}^{(I,q)}\right| + \bar{d}_z^{(I)}\right). \tag{18}$$

By setting

$$\bar{\eta}_\Delta^{(I,q)} = \bar{\eta}^{(I)}(x^{(I)}, u^{(I)}, C_z^{(I)}z^{(I)}, t) - \bar{\eta}^{(I)}(\hat{x}_H^{(I,q)}, u^{(I)}, y_{z_H}^{(I)}, t) \tag{19}$$

and taking into account Assumption 3, we obtain

$$\left|\eta^{(I)}(x^{(I)}, u^{(I)}, C_z^{(I)}z^{(I)}, t)\right| \leq \bar{\eta}^{(I)}(\hat{x}_H^{(I,q)}, u^{(I)}, y_{z_H}^{(I)}, t) + \left|\bar{\eta}_\Delta^{(I,q)}\right|, \tag{20}$$

Given that $\bar{\eta}^{(I)}$ is locally Lipschitz, we have

$$\left|\bar{\eta}_\Delta^{(I,q)}\right| \leq \lambda_{\eta_I}\left(\left|\varepsilon_{x_H}^{(I,q)}\right| + \bar{d}_z^{(I)}\right) \tag{21}$$

Based on (17)-(21), the bound on $\left|\varepsilon_{x_H}^{(I,q)}(t)\right|$ can be computed as

$$\left|\varepsilon_{x_H}^{(I,q)}(t)\right| \leq E_H^{(I,q)}(t) + \int_0^t \Lambda_I\Phi(t-\tau)\left|\varepsilon_{x_H}^{(I,q)}(\tau)\right|d\tau \tag{22}$$

where $E_H^{(I,q)}$ is defined in (15). Applying the Bellman-Gronwall Lemma [22] results in (14). ■

It is noted that in the absence of modeling uncertainty and noise i.e. assuming that $\bar{\eta}^{(I)} = 0$ and $\bar{d}^{(I,q)} = 0$, $\bar{d}_z^{(I)} = 0$, we obtain $E_H^{(I,q)}(t) = \rho^{(I,q)}e^{-\xi^{(I,q)}t}\bar{x} \to 0$. Consequently, $|\varepsilon_{x_H}^{(I,q)}(t)|$ converges to zero as $t \to \infty$. The residual under healthy conditions can be expressed as a function of the state estimation error under healthy conditions $\varepsilon_{x_H}^{(I,q)}(t)$, i.e.

$$\varepsilon_{y_{jH}}^{(I,q)} = C_j^{(I)}\varepsilon_{x_H}^{(I,q)} + d_j^{(I)}, \quad j \in \mathcal{J}^{(I,q)} \tag{23}$$

with $\varepsilon_{x_H}^{(I,q)}$ described by (11)-(13). If there is no sensor noise and modeling uncertainty, $\varepsilon_{y_{jH}}^{(I,q)}$ converges to zero as $t \to \infty$.

B. Computation of adaptive thresholds

The j-th adaptive threshold, denoted by $\bar{\varepsilon}_{y_j}^{(I,q)}(t)$, $j \in \mathcal{J}^{(I,q)}$, is designed to bound the residual under healthy conditions $\varepsilon_{y_{jH}}^{(I,q)}$. Using the solution of (11), we compute the j-th adaptive threshold following the same procedure presented in the proof of Lemma 4.1. It is noted that for the computation

of the j-th adaptive threshold, we choose positive constants $\alpha_j^{(I,q)}$, $\zeta_j^{(I,q)}$ such that $|C_j^{(I,q)} e^{A_L^{(I,q)} t}| \leq \alpha_j^{(I,q)} e^{-\zeta_j^{(I,q)} t}$ and $\alpha_{d_j}^{(I,q)}$, $\zeta_{d_j}^{(I,q)}$ such that $|C_j^{(I)} e^{A_L^{(I,q)} t} L^{(I,q)}| \leq \alpha_{d_j}^{(I,q)} e^{-\zeta_{d_j}^{(I,q)} t}$. The j-th adaptive threshold, $j \in \mathcal{J}^{(I,q)}$ is described by

$$\bar{\varepsilon}_{y_j}^{(I,q)}(t) = Y_j^{(I,q)}(t) + \int_0^t \alpha_j^{(I,q)} e^{-\zeta_j^{(I,q)}(t-\tau)} \Big(\Lambda_I Z^{(I,q)}(\tau) + \bar{\eta}^{(I)}(\hat{x}^{(I,q)}(\tau), u^{(I)}(\tau), y_z^{(I)}(\tau), \tau) \Big) d\tau, \quad (24)$$

with

$$Y_j^{(I,q)}(t) = \alpha_j^{(I,q)} \bar{x}^{(I)} e^{-\zeta_j^{(I,q)} t} + \frac{\alpha_{d_j}^{(I,q)} \bar{d}^{(I,q)}}{\zeta_{d_j}^{(I,q)}} \left(1 - e^{-\zeta_{d_j}^{(I,q)} t} \right) + \frac{\alpha_j^{(I,q)} (\lambda_{h_I} + \lambda_{\eta_I}) \bar{d}_z^{(I)}}{\zeta_j^{(I,q)}} \left(1 - e^{-\zeta_j^{(I,q)} t} \right) + \bar{d}_j^{(I)}, \quad (25)$$

$$Z^{(I,q)}(t) = E^{(I,q)}(t) + \rho^{(I,q)} \Lambda_I \int_0^t E^{(I,q)}(\tau) e^{-\nu^{(I,q)}(t-\tau)} d\tau \quad (26)$$

where $E^{(I,q)}$ is defined through (15) after replacing $\hat{x}_H^{(I,q)}$ and $y_{z_H}^{(I)}$ with $\hat{x}^{(I,q)}$ and $y_z^{(I)}$, respectively, and $\nu^{(I,q)} = \xi^{(I,q)} - \rho^{(I,q)} \Lambda_I$ is positive according to Lemma 4.1, making the adaptive threshold finite for all t. It is important to note that the j-th adaptive threshold can be implemented using linear filters [23]; i.e.,

$$\bar{\varepsilon}_{y_j}^{(I,q)}(t) = W(s) \Big[\bar{\eta}(\hat{x}^{(I,q)}(t), u^{(I)}(t), y_z^{(I)}(t), t) + \Lambda_I Z^{(I,q)}(t) \Big] + Y_j^{(I,q)}(t), \quad (27)$$

$$W(s) = \frac{\alpha_j^{(I,q)}}{s + \zeta_j^{(I,q)}}. \quad (28)$$

The notation $W(s)[z(t)]$ denotes the output of the filter $W(s)$ defined in Laplace domain with $z(t)$ as input, for any signal $z(t)$. Similarly, the signals $Z^{(I,q)}$ and $E^{(I,q)}$ can be implemented using linear filtering techniques. Considering that there is no sensor fault in the local sensor set $\mathcal{S}^{(I,q)}$ and the transmitted sensor information, let us denote the adaptive threshold under healthy conditions by $\bar{\varepsilon}_{y_{jH}}^{(I,q)}(t)$, expressed as

$$\bar{\varepsilon}_{y_{jH}}^{(I,q)}(t) = Y_j^{(I,q)}(t) + \int_0^t \alpha_j^{(I,q)} e^{-\zeta_j^{(I,q)}(t-\tau)} \Big(\Lambda_I Z_H^{(I,q)}(\tau) + \bar{\eta}^{(I)}(\hat{x}_H^{(I,q)}(\tau), u^{(I)}(\tau), y_{z_H}^{(I)}(\tau), \tau) \Big) d\tau \quad (29)$$

where $Z_H^{(I,q)}$ is determined through (26) with $E^{(I,q)} = E_H^{(I,q)}$ ($E_H^{(I,q)}$ is defined in (15)) . Hence, under healthy conditions,

$$\left| \varepsilon_{y_{jH}}^{(I,q)} \right| \leq \bar{\varepsilon}_{y_{jH}}^{(I,q)}, \quad (30)$$

where $\varepsilon_{y_{jH}}^{(I,q)}$ is defined in (23).

C. Distributed sensor fault detection decision logic

The sensor fault detection decision logic implemented in the module $\mathcal{M}^{(I,q)}$ is based on a set of analytical redundancy relations (ARRs), which are dynamical constraints, formulated using the residuals and adaptive thresholds [2], [24], [25]. Specifically, the j-th ARR, associated with the module $\mathcal{M}^{(I,q)}$ is defined as:

$$\mathcal{E}_j^{(I,q)} : \qquad \left| \varepsilon_{y_j}^{(I,q)}(t) \right| - \bar{\varepsilon}_{y_j}^{(I,q)}(t) \leq 0, \quad (31)$$

When inequality in (31) is true, it is inferred that the ARR $\mathcal{E}_j^{(I,q)}$ is satisfied. The set of ARRs, based on which the module $\mathcal{M}^{(I,q)}$ obtains a decision, is defined as

$$\mathcal{E}^{(I,q)} = \bigcup_{j \in \mathcal{J}^{(I,q)}} \mathcal{E}_j^{(I,q)}. \quad (32)$$

Therefore, the set $\mathcal{E}^{(I,q)}$ is satisfied when $\mathcal{E}_j^{(I,q)}$ is satisfied for all $j \in \mathcal{J}^{(I,q)}$. The distributed sensor fault detection decision logic is formulated taking into account the robustness and structured fault sensitivity of the set $\mathcal{E}^{(I,q)}$, which are described in the following lemma.

Lemma 4.2: Taking into account the set of ARRs $\mathcal{E}^{(I,q)}$, defined in (32), it is ensured that:

(a) **Robustness:** If neither the local sensor set $\mathcal{S}^{(I,q)}$ nor the transmitted sensor information $y_z^{(I)}$ are affected by sensor faults, then the set of ARRs $\mathcal{E}^{(I,q)}$ is always satisfied.

(b) **Structured fault sensitivity**: If there is a time instant at which $\mathcal{E}^{(I,q)}$ is not satisfied, then the occurrence of at least one sensor fault in $\mathcal{S}^{(I,q)} \bigcup \mathcal{S}_z^{(I)}$ is guaranteed.

Proof: (a) If both $\mathcal{S}^{(I,q)}$ and $\mathcal{S}_z^{(I)}$ are healthy, the j-th residual is described by (23); i.e., $\varepsilon_{y_j}^{(I,q)} = \varepsilon_{y_{jH}}^{(I,q)}$, and the j-th adaptive threshold is defined by (29), i.e., $\bar{\varepsilon}_{y_j}^{(I,q)} = \bar{\varepsilon}_{y_{jH}}^{(I,q)}$, implying that (30) is valid and consequently $\mathcal{E}^{(I,q)}$ is satisfied.

(b) The second part of Lemma 4.2 can be proved by *reductio ad absurdum*. Suppose that no sensor fault has occurred in $\mathcal{S}^{(I,q)}$ and $\mathcal{S}_z^{(I)}$. Then, $\mathcal{E}^{(I,q)}$ is satisfied, according to part (a) of Lemma 4.2. This contradicts our assumption of part (b). ■

It is noted that the robustness and structured fault sensitivity properties result from the design of the nonlinear observer, the residuals and the adaptive thresholds. The robustness property implies that the set of ARRs $\mathcal{E}^{(I,q)}$ is insensitive to modeling uncertainties and noise, thus avoiding false alarms. On the other hand, the structured fault sensitivity property entails that $\mathcal{E}^{(I,q)}$ is sensitive to a *subset* of all possible sensor faults that may affect the sensor set $\mathcal{S}^{(I)}$ (since $\mathcal{S}^{(I,q)} \subset \mathcal{S}^{(I)}$) and the sensor sets monitoring the neighboring subsystems (since $\mathcal{S}_z^{(I)} \subset \bigcup_{Q \in \mathcal{N}^{(I)}} \mathcal{S}^Q$, where $\mathcal{N}^{(I)}$ is the set including the subsystems that are physically interconnected with $\Sigma^{(I)}$).

The output of $\mathcal{M}^{(I,q)}$, denoted by $D^{(I,q)}$, is the decision on the presence of sensor faults in $\mathcal{S}^{(I,q)} \bigcup \mathcal{S}_z^{(I)}$, represented by a boolean function, defined as

$$D^{(I,q)}(t) = \begin{cases} 0, & \text{for } t < T_D^{(I,q)} \\ 1, & \text{otherwise} \end{cases}, \quad (33)$$

where $T_D^{(I,q)}$ is the detection time for the module $\mathcal{M}^{(I,q)}$, defined as

$$T_D^{(I,q)} = \min_t \bigcup_{j \in \mathcal{J}^{(I,q)}} \left\{ \min_t \left\{ t : |\varepsilon_{y_j}^{(I,q)}(t)| > \bar{\varepsilon}_{y_j}^{(I,q)}(t) \right\} \right\} \quad (34)$$

If $\mathcal{E}^{(I,q)}$ is always satisfied, then the detection time is defined as $T_{FD}^{(I,q)} = \infty$. When $D^{(I,q)}(t) = 1$, the set $\mathcal{S}^{(I,q)} \bigcup \mathcal{S}_z^{(I)}$ is faulty, since at least one sensor fault in $\mathcal{S}^{(I,q)} \bigcup \mathcal{S}_z^{(I)}$ is guaranteed to have occurred, according to the structured fault sensitivity (Lemma 4.2). As long as $D^{(I,q)}(t) = 0$ either there is no sensor fault in $\mathcal{S}^{(I,q)} \bigcup \mathcal{S}_z^{(I)}$ or sensor faults have occurred, but have not been detected by the module $\mathcal{M}^{(I,q)}$ until the time $T_{FD}^{(I,q)}$. However, the sensor set $\mathcal{S}^{(I,q)} \bigcup \mathcal{S}_z^{(I)}$ is characterized as *non-faulty*, based on the exoneration assumption [17], [24]; i.e., given a set of observations, the sensors in $\mathcal{S}^{(I,q)} \bigcup \mathcal{S}_z^{(I)}$ necessarily reveal their faulty operation by provoking the violation of $\mathcal{E}^{(I,q)}$, or equivalently, all sensors in $\mathcal{S}^{(I,q)} \bigcup \mathcal{S}_z^{(I)}$ are exonerated, i.e. are considered as functioning properly, if $\mathcal{E}^{(I,q)}$ is satisfied.

V. Distributed Sensor Fault Isolation

The multiple sensor fault isolation is realized in two levels; locally, by combining the decisions of the monitoring modules and globally, by combining the decisions of the monitoring agents of the relevant CPS.

A. Multiple sensor fault isolation decision logic of CPS

The monitoring agent $\mathcal{M}^{(I)}$ uses a binary fault signature matrix $F^{(I)}$, consisting of N_I rows and $Nc_I + 2$ columns, where $Nc_I = 2^{m_I} - 1$ (m_I is the number of sensors in the sensor set $\mathcal{S}^{(I)}$); the q-th row corresponds to the q-th set of ARRs $\mathcal{E}^{(I,q)}$, $q \in \{1, \ldots, N_I\}$; the i-th column, for all $i \in \{1, \ldots, Nc_I\}$ corresponds to the i-th combination of sensor faults that may affect the sensor set $\mathcal{S}^{(I)}$, denoted by $\mathcal{F}_{c_i}^{(I)}$, while the $Nc_I + 1$ column corresponds to the fault vector $f_z^{(I)}$, that is sensor faults from neighboring CPS, and the $Nc_I + 2$ column corresponds to the union of all combinations of $f_z^{(I)}$ and $\mathcal{F}_{c_i}^{(I)}$ for all $i \in \{1, \ldots, Nc_I\}$. For example, if $m_I = 2$, $F^{(I)}$ has five columns, corresponding to the following sensor fault combination: $\mathcal{F}_{c_1}^{(I)} = \left\{f_1^{(I)}\right\}$, $\mathcal{F}_{c_2}^{(I)} = \left\{f_2^{(I)}\right\}$, $\mathcal{F}_{c_3}^{(I)} = \left\{f_1^{(I)}, f_2^{(I)}\right\}$, $\mathcal{F}_{c_4}^{(I)} = \left\{f_z^{(I)}\right\}$ and $\mathcal{F}_{c_5}^{(I)} = \bigcup_{i \in \{1,2,3\}} \left\{f_z^{(I)}, \mathcal{F}_{c_i}^{(I)}\right\}$. The i-th column, denoted by $F_i^{(I)}$, corresponds to the theoretical pattern of sensor faults defined as:

$$F_i^{(I)} = \left[F_{1i}^{(I)}, \ldots, F_{N_I i}^{(I)}\right]^\top, \tag{35}$$

where $F_{qi}^{(I)} = 1$, if at least one sensor fault included in the combination $\mathcal{F}_{c_i}^{(I)}$, $i \in \{1, \ldots, Nc_I\}$, can provoke the violation of (or else is involved in) $\mathcal{E}^{(I,q)}$, $q \in \{1, \ldots, N_I\}$ and $F_{qi}^{(I)} = 0$ otherwise (see simulation example in Section VII). The design of $F^{(I)}$ exploits the structure sensor fault sensitivity property of $\mathcal{E}^{(I,q)}$, described in Lemma 4.2.

The decisions obtained by the N_I modules of the agent $\mathcal{M}^{(I)}$ constitute the observed pattern of sensor faults affecting $\mathcal{S}^{(I)}$ and $\mathcal{S}_z^{(I)}$, denoted by $D^{(I)}(t)$; i.e.,

$$D^{(I)}(t) = \left[D^{(I,1)}(t), \ldots, D^{(I,N_I)}(t)\right]^\top \tag{36}$$

where $D^{(I,q)}$, $q \in \{1, \ldots, N_I\}$ is defined in (33). The observed pattern, $D^{(I)}(t)$ is compared to each of the $Nc_I + 2$ theoretical patterns $F_i^{(I)}$, $I \in \{1, \ldots, N\}$, in order to determine the sensor fault diagnosis set $\mathcal{D}_s^{(I)}(t)$ that includes the sensor fault combinations that have possibly occurred [26]. As long as $D^{(I)}(t) = \mathbf{0}_{N_I}$ ($\mathbf{0}_{N_I}$ is a zero vector of length N_I), the diagnosis set $\mathcal{D}_s^{(I)}(t)$ is empty; otherwise, if $D^{(I,q)}(t) = F_{qi}^{(I)}$ for all $q \in \{1, \ldots, N_I\}$, then the observed pattern $D^{(I)}(t)$ is said to be consistent with the i-th theoretical pattern $F_i^{(I)}$ (*consistency test*) and the diagnosis set is defined as

$$\mathcal{D}_s^{(I)}(t) = \left\{\mathcal{F}_{c_i}^{(I)} : i \in \mathcal{I}_{\mathcal{D}}^{(I)}(t)\right\}, \tag{37}$$

where $\mathcal{I}_{\mathcal{D}}^{(I)}(t)$ is the consistency index set defined as $\mathcal{I}_{\mathcal{D}}^{(I)}(t) = \left\{i : F_i^{(I)} = D^{(I)}(t),\ i \in \{1, \ldots, Nc_I\}\right\}$. The diagnosis set $\mathcal{D}_s^{(I)}(t)$ may contain one or more fault combinations.

The observed pattern $D^{(I)}(t)$ changes over time, thus it is possible that at some time instant the consistency test is not satisfied. This may happen when there are two or more identical theoretical patterns and some possible observed patterns cannot be consistent with the theoretical patterns. When the consistency test does not provide any result, the sensor fault diagnosis set $\mathcal{D}_s^{(I)}(t)$ contains the sensor fault combinations involved in the violated ARRs [24]; i.e.

$$\mathcal{D}_s^{(I)}(t) = \bigcap_{q \in \mathcal{Q}(t)} Supp(\mathcal{E}^{(I,q)}), \tag{38}$$

where $Supp(\mathcal{E}^{(I,q)})$ is the support of $\mathcal{E}^{(I,q)}$, which is the set of sensor fault combinations $\mathcal{F}_{c_i}^{(I)}$ for which $F_{qi}^{(I)} = 1$ and $\mathcal{Q}^{(I)}(t)$ is the index set of the violated ARRs, defined as $\mathcal{Q}^{(I)}(t) = \left\{q : D^{(I,q)}(t) = 1,\ q \in \{1, \ldots, N_I\}\right\}$. In general, $\left\{\mathcal{F}_{c_i}^{(I)} : i \in \mathcal{I}_{\mathcal{D}}^{(I)}(t)\right\} \subseteq \bigcap_{q \in \mathcal{Q}(t)} Supp(\mathcal{E}^{(I,q)})$, i.e. the consistency test provides a diagnosis set with a smaller (or equal) number of diagnosed sensor fault combinations compared to the diagnosis set defined in (38).

The outputs of the agent $\mathcal{M}^{(I)}$, $I \in \{1, \ldots, N\}$ are the diagnosis set $\mathcal{D}_s^{(I)}$ and the decision on the presence of sensor faults in $\mathcal{S}_z^{(I)}$, represented by the function $D_z^{(I)}(t)$:

$$D_z^{(I)}(t) = \begin{cases} 0, & \text{if } f_z^{(I)} \notin \mathcal{D}_s^{(I)}(t) \\ 1, & \text{if } f_z^{(I)} \in \mathcal{D}_s^{(I)}(t) \end{cases} \tag{39}$$

The diagnostic reasoning behind the decision of the agent $\mathcal{M}^{(I)}$ on multiple sensor fault isolation relies on the resultant diagnosis set. Particularly, we may infer that (i) at least one of the sensor fault combinations $\mathcal{F}_{c_i}^{(I)}$ that belongs to $\mathcal{D}_s^{(I)}(t)$ has occurred, (ii) the sensor fault $f_j^{(I)}$ included in all diagnosed sensor fault combinations, i.e., $f_j^{(I)} \in \bigcap_{i \in \mathcal{I}_{\mathcal{D}}^{(I)}(t)} \mathcal{F}_{c_i}^{(I)}$, with $\bigcap_{i \in \mathcal{I}_{\mathcal{D}}^{(I)}(t)} \mathcal{F}_{c_i}^{(I)} \neq \emptyset$, is guaranteed to have occurred, and (iii) the occurrence of the combination $\mathcal{F}_{c_i}^{(I)}$ that does not belong to the set of sensor fault combinations involved in the violated $\mathcal{E}^{(I,q)}$, i.e. $\mathcal{F}_{c_i}^{(I)} \notin \bigcap_{q \in \mathcal{Q}(t)} Supp(\mathcal{E}^{(I,q)})$, is excluded. When $f_j^{(I)}$ is guaranteed to have occurred, the sensor $\mathcal{S}^{(I)}\{j\}$

is isolated as faulty. If $f_j^{(I)}$ belongs to some of the diagnosed sensor fault combinations, but not to all of them, the sensor $\mathcal{S}^{(I)}\{j\}$ is characterized as possibly faulty. If $f_z^{(I)}$ belongs to the diagnosis set, then at least one sensor fault in $\mathcal{S}_z^{(I)}$ may have been propagated to $\mathcal{M}^{(I)}$ from a neighboring agent.

B. Global sensor fault isolation decision logic

The primary goal of the global decision logic $\mathcal{G}$ is to isolate sensor faults that have been propagated from neighboring agents. If $D_z^{(I)}(t) = 0$ for all $I \in \{1, \ldots, N\}$, then no information (decisions) is processed by the global decision logic; otherwise the global decision logic is based on the theoretical patterns of propagated sensor faults, which describe the involvement of the sensor faults $f_z^{(I)}$ in the set of ARRs, $\mathcal{E}^{(I)}$, for all $I \in \{1, \ldots, N\}$, defined as

$$\mathcal{E}^{(I)} = \bigcup_{q \in \{1,\ldots,N_I\}} \mathcal{E}^{(I,q)}. \tag{40}$$

The theoretical patterns are the columns of the sensor fault signature matrix F^z, which has N rows and $Nc = 2^p - 1$ columns, where $p \leq \sum_{I=1}^{N} p_I$ (p_I is the length of $f_z^{(I)}$). The I-th row corresponds to the set of ARRs $\mathcal{E}^{(I)}$ and the k-th column, $k \in \{1, \ldots, Nc\}$ to the k-th combination of sensor faults, denoted by $\mathcal{F}_{c_k}^z$ that affect $\bigcup_{I=1}^{N} \mathcal{S}_z^{(I)}$. The k-th theoretical pattern of transmitted faulty sensor information is the k-th column of F^z defined as:

$$F_k^z = [F_{1k}^z, \ldots, F_{Nk}^z]^\top \tag{41}$$

where: a) $F_{Ik}^z = 1$, $I \in \{1, \ldots, N\}$, $k \in \{1, \ldots, Nc\}$, if at least one sensor fault included in $\mathcal{F}_{c_k}^z$ is involved in $\mathcal{E}^{(I)}$ and affects $\mathcal{S}^{(I)}$, b) $F_{Ik}^z = 0$ if none of the sensor faults included in $\mathcal{F}_{c_k}^z$ is involved in $\mathcal{E}^{(I)}$, and c) $F_{Ik}^z = *$, if at least one sensor fault included in $\mathcal{F}_{c_k}^z$ is involved in $\mathcal{E}^{(I)}$ and affects $\mathcal{S}_z^{(I)}$, while none of the sensor faults included in $\mathcal{F}_{c_k}^z$ affects $\mathcal{S}^{(I)}$ (see simulation example in Section VII).

The semantics of $F_{Ik}^z = 1$ is that a sensor fault of $\mathcal{F}_{c_k}^z$ that affects $\mathcal{S}^{(I)}$ necessarily provokes the violation of $\mathcal{E}^{(I)}$. In other words, $\mathcal{E}^{(I)}$ is very sensitive to sensor faults $f^{(I)}$. The semantics of $F_{Ik}^z = *$ is that a sensor fault of $\mathcal{F}_{c_k}^z$ that belongs to $f_z^{(I)}$ and is involved in $\mathcal{E}^{(I)}$, i.e. a sensor fault that is propagated from a neighboring agent, can explain why $\mathcal{E}^{(I)}$ is violated at some time instant, but $\mathcal{E}^{(I)}$ may happen to be satisfied while this sensor fault in $\mathcal{F}_{c_k}^z$ has occurred. This may happen if $\mathcal{E}^{(I)}$ is not very sensitive to $f_z^{(I)}$ (see Section VI).

The combinatorial process of the decisions of the N monitoring agents is performed by initially determining the observed pattern of propagated sensor faults as:

$$D_z(t) = \left[D_z^{(1)}(t), \ldots, D_z^{(N)}(t)\right]^\top. \tag{42}$$

The observed pattern $D_z(t)$ is compared to each of the columns of F_k^z. If $D_z(t)$ is consistent to F_k^z, i.e. $D_z(t) = F_k^{(z)}$ the diagnosis set of propagated sensor faults is defined as:

$$\mathcal{D}_s^z(t) = \left\{\mathcal{F}_{c_k}^z : k \in \mathcal{I}_z(t)\right\}, \tag{43}$$

where $\mathcal{I}_z(t)$ is an index set defined as $\mathcal{I}_z(t) = \left\{k : D_z^{(I)}(t) = F_{Ik}^z, \forall\, I \in \{1, \ldots, N\}, k \in \{1, \ldots, Nc\}\right\}$. If $F_{Ik}^z = *$ then $D_z^{(I)}(t) = F_{Ik}^z$ if either $D_z^{(I)}(t) = 0$ or $D_z^{(I)}(t) = 1$. If $D_z(t)$ is not consistent to F_k^z, then $\mathcal{D}_s^z(t)$ contains the sensor fault combinations that belong to the support of the violated $\mathcal{E}^{(I)}$, i.e. the sensor fault combinations $\mathcal{F}_{c_k}^z$ for which $F_{Ik}^z = 1$ or $F_{Ik}^z = *$. The set $\mathcal{D}_s^z(t)$ is used to update the non-empty diagnosis set $\mathcal{D}_s^{(I)}(t)$ of $\mathcal{M}^{(I)}$ by excluding $f_z^{(I)}$ and its combinations, if $f_z^{(I)} \notin \mathcal{D}_s^z(t)$.

VI. Performance Analysis

In this section, we analyze the performance of the proposed distributed SFDI scheme with respect to the distributed sensor fault effects and fault detectability of the modules $\mathcal{M}^{(I,q)}$.

A. Distributed Sensor Fault Effects

The effects of sensor faults on the j-th residual and adaptive threshold, denoted by $\varepsilon_{y_{jF}}^{(I,q)}$ and $\bar{\varepsilon}_{y_{jF}}^{(I,q)}$, $j \in \mathcal{J}^{(I,q)}$, respectively, for $t \geq T^*$, where T^* is the first time instant of sensor fault occurrence, can be determined as

$$\varepsilon_{y_{jF}}^{(I,q)} = \varepsilon_{y_j}^{(I,q)} - \varepsilon_{y_{jH}}^{(I,q)}, \tag{44}$$

$$\bar{\varepsilon}_{y_{jF}}^{(I,q)} = \bar{\varepsilon}_{y_j}^{(I,q)} - \bar{\varepsilon}_{y_{jH}}^{(I,q)}, \tag{45}$$

where $\varepsilon_{y_{jH}}^{(I,q)}$, $\bar{\varepsilon}_{y_{jH}}^{(I,q)}$ are defined in (23) and (29), respectively.

For $t \geq T^*$, the estimation model of $\mathcal{M}^{(I,q)}$ satisfies

$$\begin{aligned}\dot{\hat{x}}^{(I,q)} = & A_L^{(I,q)}\hat{x}^{(I,q)} + \gamma^{(I)}(\hat{x}^{(I,q)}, u^{(I)}) \\ & + h^{(I)}(\hat{x}^{(I,q)}, u^{(I)}, y_{z_H}^{(I)} + f_z^{(I)}) \\ & + L^{(I,q)}\left(y_H^{(I,q)} + f^{(I,q)}\right),\end{aligned} \tag{46}$$

where $y_H^{(I,q)}$, $y_{z_H}^{(I)}$ are defined in (7) and (6) with $f^{(I,q)} = 0$ and $f_z^{(I)} = 0$. Given that $\varepsilon_{y_j}^{(I,q)} = y_{jH}^{(I)} + f_j^{(I)} - C_j^{(I)}\hat{x}^{(I,q)}$ and $\varepsilon_{y_{jH}}^{(I,q)} = y_{jH}^{(I)} - C_j^{(I)}\hat{x}_H^{(I,q)}$, $j \in \mathcal{J}^{(I,q)}$, the effects $\varepsilon_{y_{jF}}^{(I,q)}(t)$ for $t \geq T^*$ can be described by

$$\varepsilon_{y_{jF}}^{(I,q)} = -C_j^{(I)}\left(\hat{x}^{(I,q)} - \hat{x}_H^{(I,q)}\right) + f_j^{(I)}, \tag{47}$$

where $\hat{x}^{(I,q)}$ is described by (46). Based on (10) and (46),

$$\begin{aligned}\hat{x}^{(I,q)}(t) - \hat{x}_H^{(I,q)}(t) = \int_{T^*}^{t} e^{A_L^{(I,q)}(t-\tau)} \Big(& \gamma^{(I)}(\hat{x}^{(I,q)}(\tau), u^{(I)}(\tau)) \\ & - \gamma^{(I)}(\hat{x}_H^{(I,q)}(\tau), u^{(I)}(\tau)) \\ & + h^{(I)}(\hat{x}^{(I,q)}(\tau), u^{(I)}(\tau), y_{z_H}^{(I)}(\tau) + f_z^{(I)}(\tau)) \\ & - h^{(I)}(\hat{x}_H^{(I,q)}(\tau), u^{(I)}(\tau), y_{z_H}^{(I)}(\tau)) \\ & + L^{(I,q)} f^{(I,q)}(\tau)\Big)\end{aligned} \tag{48}$$

If $|f_z^{(I)}| < \infty$ and $|f^{(I,q)}| < \infty$, a bound on $|\hat{x}^{(I,q)} - \hat{x}_H^{(I,q)}|$ denoted by $\bar{\varepsilon}_{x_F}^{(I,q)}$, can be derived using Assumption 2 and the Bellman-Gronwall Lemma [22]; i.e.,

$$\begin{aligned}\bar{\varepsilon}_{x_F}^{(I,q)} = & \int_{T^*}^{t} \rho^{(I,q)}\left(\lambda_{\gamma_I} + \lambda_{h_I}\right) e^{-\mu^{(I,q)}(t-\tau)} \Upsilon_z^{(I,q)}(\tau) d\tau \\ & + \Upsilon_z^{(I,q)}(t),\end{aligned} \tag{49}$$

$$\Upsilon_z^{(I,q)}(t) = \lambda_{h_I} \int_{T^*}^{t} \rho^{(I,q)} e^{-\xi^{(I,q)}(t-\tau)} \left| f_z^{(I)}(\tau) \right| d\tau + |L^{(I,q)}| \int_{T^*}^{t} \rho^{(I,q)} e^{-\xi^{(I,q)}(t-\tau)} \left| f^{(I,q)}(\tau) \right| d\tau \quad (50)$$

where $\mu^{(I,q)} = \xi^{(I,q)} - \rho^{(I,q)}(\lambda_{\gamma_I} + \lambda_{h_I})$ is positive. By introducing (50) in (49), we obtain a compact form for $\bar{\varepsilon}_{x_F}^{(I,q)}$

$$\bar{\varepsilon}_{x_F}^{(I,q)} = \lambda_{h_I} g_j^{(I,q)}(|f_z^{(I)}|) + |L^{(I,q)}| g_j^{(I,q)}(|f^{(I,q)}|), \quad (51)$$

where $g_j^{(I,q)}$ represents the propagated and local sensor fault effects on the j-th residual of the module $\mathcal{M}^{(I,q)}$, $j \in \mathcal{J}^{(I,q)}$. The effects $\bar{\varepsilon}_{y_{jF}}^{(I,q)}(t)$ can be expressed as

$$\bar{\varepsilon}_{y_{jF}}^{(I,q)}(t) = \int_{T^*}^{t} \alpha_j^{(I,q)} e^{-\zeta_j^{(I,q)}(t-\tau)} \left(\delta\bar{\eta}_F^{(I,q)}(\tau) + \Lambda_I \delta Z^{(I,q)}(\tau) \right) d\tau, \quad (52)$$

$$\delta\bar{\eta}_F^{(I,q)} = \bar{\eta}^{(I)}(\hat{x}^{(I,q)}, u^{(I)}, y_{z_H}^{(I)} + f_z^{(I)}) - \bar{\eta}^{(I)}(\hat{x}_H^{(I,q)}, u^{(I)}, y_{z_H}^{(I)}), \quad (53)$$

where $\delta Z^{(I,q)} = Z^{(I,q)} - Z_H^{(I,q)}$, defined through (26) by replacing $E^{(I,q)}$ with $\delta E^{(I,q)} = E^{(I,q)} - E_H^{(I,q)}$, defined as

$$\delta E^{(I,q)} = \int_{T^*}^{t} \rho^{(I,q)} e^{-\xi^{(I,q)}(t-\tau)} \delta\bar{\eta}_F^{(I,q)}(\tau)\, d\tau. \quad (54)$$

Using Assumption 3 and $|\hat{x}^{(I,q)} - \hat{x}_H^{(I,q)}| \leq \bar{\varepsilon}_{x_F}^{(I,q)}$, where $\bar{\varepsilon}_{x_F}^{(I,q)}$ is defined in (51), we obtain

$$\left| \delta\bar{\eta}_F^{(I,q)} \right| \leq \lambda_{\eta_I} \lambda_{h_I} g_j^{(I,q)}(|f_z^{(I)}|) + \lambda_{\eta_I} |f_z^{(I)}| + \lambda_{\eta_I} |L^{(I,q)}| g_j^{(I,q)}(|f^{(I,q)}|). \quad (55)$$

A bound on $\delta E^{(I,q)}$ can be defined as

$$\left| \delta E^{(I,q)} \right| \leq \int_{T^*}^{t} \rho^{(I,q)} e^{-\xi^{(I,q)}(t-\tau)} \left(\lambda_{\eta_I} \left(\lambda_{h_I} g_j^{(I,q)}(|f_z^{(I)}(\tau)|) \right) + \lambda_{\eta_I} \left| f_z^{(I)}(\tau) \right| + \lambda_{\eta_I} |L^{(I,q)}| g_j^{(I,q)}(|f^{(I,q)}(\tau)|) \right) d\tau \quad (56)$$

Taking into account (26), a bound on $\delta Z^{(I,q)}$ is determined as

$$\left| \delta Z^{(I,q)} \right| \leq \left| \delta E^{(I,q)} \right| + \rho^{(I,q)} \Lambda_I \int_0^t e^{-\nu^{(I,q)}(t-\tau)} \left| \delta E^{(I,q)}(\tau) \right| d\tau \quad (57)$$

where $\nu^{(I,q)} = \xi^{(I,q)} - \rho^{(I,q)}\Lambda_I$ is positive (see Lemma 4.1). Combining (55)-(57), a bound on $\left| \bar{\varepsilon}_{y_{jF}}^{(I,q)} \right|$ is defined as

$$\left| \bar{\varepsilon}_{y_{jF}}^{(I,q)} \right| \leq \lambda_{h_I} \bar{g}_{Aj}^{(I,q)}(|f_z^{(I)}|) + \bar{g}_{Bj}^{(I,q)}(|f_z^{(I)}|) + |L^{(I,q)}| \bar{g}_{Aj}^{(I,q)}(|f^{(I,q)}|) \quad (58)$$

where $g_{Aj}^{(I,q)}$, $\bar{g}_{Bj}^{(I,q)}$ are functions describing an upper bound for the sensor fault effects on the j-th adaptive threshold of the module $\mathcal{M}^{(I,q)}$. Taking into account (58) and that

$$\left| \varepsilon_{y_{jF}}^{(I,q)} \right| \leq \lambda_{h_I} \left| C_j^{(I)} \right| g_j^{(I,q)}(|f_z^{(I)}|) + \left| L^{(I,q)} \right| \left| C_j^{(I)} \right| g_j^{(I,q)}(|f^{(I,q)}|) + \left| f_j^{(I)} \right|, \quad (59)$$

we may infer that the sensor fault propagation effects on the j-th residual and adaptive threshold depend on the interconnection function, represented by the Lipschitz constant λ_{h_I}. This means that propagated sensor faults of large magnitude may have low impact on the j-th residual and adaptive threshold or propagated sensor faults of small magnitude may have high impact on the j-th residual and adaptive threshold. On the other hand, the effects of sensor faults $f^{(I)}$ on the j-th residual and adaptive threshold can be designed to be higher than the sensor fault propagation effects, because they are amplified by the magnitude of the observer gain $|L^{(I,q)}|$. This quantitative analysis is the basis of the design of the sensor fault signature matrix implemented in $\mathcal{G}$, which may differentiate qualitatively the sensitivity of $\mathcal{E}^{(I)}$ with respect to the propagated sensor faults $f_z^{(I)}$ and the local sensor faults $f^{(I)}$ in the I-th CPS.

B. Distributed Sensor Fault Detectability

The conditions that characterize the minimum effects of sensor faults in $\mathcal{S}^{(I,q)}$ and/or propagated sensor faults that are detectable by the module $\mathcal{M}^{(I,q)}$ are given in the following Lemma.

Lemma 6.1: The occurrence of faults in $\mathcal{S}^{(I,q)}$ and/or $\mathcal{S}_z^{(I)}$ is guaranteed to be detected under worst-case conditions, if there exists a time instant $t^\circ \geq T^*$(T^* is the first time instant of sensor fault occurrence) such that the effects of sensor faults $f^{(I,q)}$ and/or $f_z^{(I)}$ on the j-th residual and adaptive threshold satisfy the condition

$$\left| \varepsilon_{y_{jF}}^{(I,q)}(t^\circ) \right| - \bar{\varepsilon}_{y_{jF}}^{(I,q)}(t^\circ) > 2\bar{\varepsilon}_{y_{jH}}^{(I,q)}(t^\circ), \quad (60)$$

where $\varepsilon_{y_{jF}}^{(I,q)}$ and $\bar{\varepsilon}_{y_{jF}}^{(I,q)}$ are defined through (44) and (45).

Proof: Due to page limitations, the proof is omitted. A similar proof can be found in [23]. ■

It is important to note that the class of detectable sensor faults satisfying (60) are obtained under worst-case detectability conditions in the sense that they are valid for any modeling uncertainty and measurement noise, given Assumptions 1-4. This condition in combination with the effects of sensor faults in $\mathcal{S}^{(I,q)}$ and propagated sensor faults can be taken into account during the design of the modules in $\mathcal{M}^{(I,q)}$, since it provides a relationship between the sensor faults $f^{(I,q)}$, $f_z^{(I)}$ and the selected design parameters used for the implementation of the nonlinear observer $\mathscr{O}^{(I,q)}$ (e.g. $L^{(I,q)}$) and the adaptive thresholds ($\rho^{(I,q)}, \xi^{(I,q)}$, $\alpha_j^{(I,q)}, \zeta_j^{(I,q)}$, $j \in \mathcal{J}^{(I,q)}$), as well as system characteristics ($\bar{\eta}^{(I)}$, $\bar{d}^{(I)}$, λ_{γ_I}, λ_{h_I}, λ_{η_I}). For instance, assuming the occurrence of sensor faults only in $\mathcal{S}_z^{(I)}$ ($f^{(I,q)} = 0$), the module $\mathcal{M}^{(I,q)}$ is guaranteed to detect them at some time instant t° if

$$2\bar{\varepsilon}_{y_{jH}}^{(I,q)}(t^\circ) < \left| \varepsilon_{y_{jF}}^{(I,q)}(t^\circ) \right| - \bar{\varepsilon}_{y_{jF}}^{(I,q)}(t^\circ)$$

&

$$\left| \varepsilon_{y_{jF}}^{(I,q)}(t^\circ) \right| - \bar{\varepsilon}_{y_{jF}}^{(I,q)}(t^\circ) \leq \left| C_j^{(I)} \right| \lambda_{h_I} g_j^{(I,q)}(|f_z^{(I)}(t^\circ)|) + \lambda_{h_I} \bar{g}_{Aj}^{(I,q)}(|f_z^{(I)}(t^\circ)|) + \bar{g}_{Bj}^{(I,q)}(|f_z^{(I)}(t^\circ)|).$$

Thus, if λ_{h_I} and λ_{η_I} are small, then the threshold $\bar{\varepsilon}_{y_{jH}}^{(I,q)}$ should be small or $|f_z^{(I)}(t^\circ)|$ large enough in order for $\mathcal{M}^{(I,q)}$ to guarantee the detection of the propagated sensor faults.

VII. Simulation Example

In this section, we illustrate the two-level diagnosis decision logic for detecting and isolating multiple sensor faults using a two-zone Heating, Ventilation and Air-Conditioning (HVAC) system (Fig. 3), which is a nonlinear system, comprised of two separated zones and the electromechanical part.

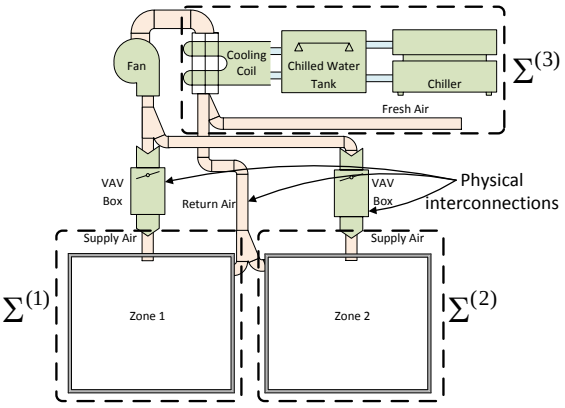


Fig. 3. Two-zone HVAC system.

Let us denote $\Sigma^{(1)}$ and $\Sigma^{(2)}$ the two subsystems that correspond to zone 1 and 2 and $\Sigma^{(3)}$ that corresponds to the electromechanical part. The temperature dynamics in each zone and the electromechanical part can be modeled based on the fundamental mass and energy conservation equations; i.e.

$$\Sigma^{(I)}: \dot{x}^{(I)}(t) = A^{(I)}x^{(I)}(t) + \gamma^{(I)}(x^{(I)}(t), u^{(I)}(t)) + h^{(I)}(z^{(I)}(t), u^{(I)}(t)) + \eta^{(I)}(t), \quad (61)$$

with $A^{(I)} = -\frac{U_{z_I}A_{z_I}}{M_{z_I}C_v}$,

$$\gamma^{(I)}(x^{(I)}, u^{(I)}) = -\frac{\rho_a C_{pa}}{M_{z_I}C_v}x^{(I)}u^{(I)} + \frac{U_{z_I}A_{z_I}}{M_{z_I}C_v}T_{amb} \quad (62)$$

$$h^{(I)}(z^{(I)}, u^{(I)}) = \frac{\rho_a C_{pa}}{M_{z_I}C_v}z^{(I)}u^{(I)} \quad (63)$$

where $x^{(I)} \in \mathbb{R}$ is the temperature of the I-th zone, $I = 1, 2$, $u^{(I)} \in \mathbb{R}$ is the volumetric flow rate of air entering into the I-th zone generated by a feedback linearization controller, and $z^{(I)} \in \mathbb{R}$ is the interconnection signal that corresponds to the temperature of output air of the cooling coil (i.e. $z^{(I)} = x_1^{(3)}$ defined below) that is transferred to the I-th zone through the physical interconnections (Fig. 3) and $\eta^{(I)}(t)$ is a disturbance signal related to the rate of internal heat gain due to occupants and appliances in the I-th zone. The subsystem $\Sigma^{(3)}$ is described by

$$\Sigma^{(3)}: \dot{x}^{(3)}(t) = A^{(3)}x^{(3)}(t) + \gamma^{(3)}(u^{(3)}(t)) + h^{(3)}(x^{(3)}(t), z^{(3)}(t), u_z^{(3)}(t)), \quad (64)$$

with

$$A^{(3)} = \begin{bmatrix} -\frac{U_{cc}A_{cc}}{M_{cc}C_v} & \frac{Q_w\rho_w C_{pw}}{M_{cc}C_v} \\ 0 & -\frac{Q_w\rho_w C_{pw}+U_tA_t}{V_t\rho_w C_{pw}} \end{bmatrix}, \quad (65)$$

$$\gamma^{(3)}(u^{(3)}) = \begin{bmatrix} \frac{U_{cc}A_{cc}}{M_{cc}C_v}T_{amb} - \frac{Q_w\rho_w C_{pw}}{M_{cc}C_v}T_{wo} \\ \frac{U_tA_t}{V_t\rho_w C_{pw}}T_{amb} + \frac{Q_w\rho_w C_{pw}}{V_t\rho_w C_{pw}}T_{wo} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{15000}{V_t\rho_w C_{pw}} \end{bmatrix} u^{(3)}, \quad (66)$$

$$h^{(3)}(x^{(3)}, z^{(3)}, u_z^{(3)}) = \begin{bmatrix} h_1^{(3)}(x^{(3)}, z^{(3)}, u_z^{(3)}) \\ 0 \end{bmatrix} \quad (67)$$

where $x^{(3)} = \left[x_1^{(3)}, x_2^{(3)}\right]^\top$ is the state vector with $x_1^{(3)}, x_2^{(3)}$ be the temperature of output air of the cooling coil and the temperature of the water in the chiller storage tank, respectively, the input $u^{(3)}$ is the chilled water mass flow rate, generated by a backstepping controller, $z^{(3)}$ is the interconnection vector whose elements are the temperature of the two zones, i.e. $z^{(3)} = [x^{(1)}; x^{(2)}]$, $u_z^{(3)}(t) = \left[u^{(1)}(t), u^{(2)}(t)\right]^\top$, and $h_1^{(3)}(x^{(3)}, z^{(3)}, u_z^{(3)})$ is defined in (68) at the bottom of this page. The parameters used for the simulation of $\Sigma^{(I)}$ $I = 1, 2$ and $\Sigma^{(3)}$ are given [27]. The temperature of each zone is measured by a sensor, denoted by $\mathcal{S}^{(I)}\{1\}$, $I = 1, 2$, while the temperature of the output air of the cooling coil and the chilled water tank are measured by two sensors, denoted by $\mathcal{S}^{(3)}\{1\}$, $\mathcal{S}^{(3)}\{2\}$, respectively.

For each of the interconnected subsystems, we design a monitoring agent $\mathcal{M}^{(I)}$, $I = 1, 2, 3$ (Fig. 4); the agents $\mathcal{M}^{(1)}$ and $\mathcal{M}^{(2)}$ monitor the sensors $\mathcal{S}^{(1)}\{1\}$ and $\mathcal{S}^{(2)}\{1\}$, respectively, using the measurements of $\mathcal{S}^{(3)}\{1\}$ as well (i.e. $y_z^{(I)} = y_1^{(3)}$, $I = 1, 2$). The agent $\mathcal{M}^{(3)}$ is comprised of two modules $\mathcal{M}^{(3,1)}$ and $\mathcal{M}^{(3,2)}$ that monitor the sensors $\mathcal{S}^{(3)}\{1\}$ and $\mathcal{S}^{(3)}\{2\}$, respectively; the observer of $\mathcal{M}^{(3,1)}$ estimates the state vector $x^{(3)}$ using the measurements of $\mathcal{S}^{(1)}\{1\}$ and $\mathcal{S}^{(2)}\{1\}$ (i.e. $y_z^{(I)} = [y_1^{(1)}; y_1^{(2)}]^\top$); the estimator of $\mathcal{M}^{(3,2)}$ is used to estimate the state $x_2^{(3)}$ and is designed taking into account the dynamic equation of $x_2^{(3)}$ described by (64)-(68) without using any transmitted sensor information.

The sensor fault signature matrix $F^{(I)}$ designed in the agent $\mathcal{M}^{(I)}$, $I = 1, 2$ for detecting the presence of faults affecting $\mathcal{S}^{(I)}\{1\}$ and/or $\mathcal{S}^{(3)}\{1\}$ is presented in the Table I. The sensor fault signature matrix $F^{(3)}$ of the agent $\mathcal{M}^{(3)}$ for isolating multiple faults affecting $\mathcal{S}^{(3)}\{1\}$, $\mathcal{S}^{(3)}\{2\}$ is presented in the Table II, with $\mathcal{F}_{c_1}^{(3)} = \left\{f_1^{(3)}\right\}$, $\mathcal{F}_{c_2}^{(3)} = \left\{f_2^{(3)}\right\}$, $\mathcal{F}_{c_3}^{(3)} = \left\{f_1^{(3)}, f_2^{(3)}\right\}$, $\mathcal{F}_{c_4}^{(3)} = \left\{f_z^{(3)}\right\}$, $\mathcal{F}_{c_5}^{(3)} = \left\{f_z^{(3)}, \mathcal{F}_{c_1}^{(3)}\right\}$, $\mathcal{F}_{c_6}^{(3)} = \left\{f_z^{(3)}, \mathcal{F}_{c_2}^{(3)}\right\}$ and $\mathcal{F}_{c_7}^{(3)} = \left\{f_z^{(3)}, \mathcal{F}_{c_3}^{(3)}\right\}$ and $f_z^{(3)} = [f_1^{(1)}, f_1^{(2)}]^\top$. The global decision logic is based on the sensor fault signature matrix F^z, shown in Table III, with $\mathcal{F}_{c_1}^z = \left\{f_1^{(1)}\right\}$, $\mathcal{F}_{c_2}^z = \left\{f_1^{(2)}\right\}$, $\mathcal{F}_{c_3}^z = \left\{f_1^{(3)}\right\}$,

$$h_1^{(3)}(x^{(3)}, z^{(3)}, u_z^{(3)}) = \left(\frac{\rho_a C_{pa}}{M_{cc}C_v}\begin{bmatrix} 1 \\ 1 \end{bmatrix}^\top u_z^{(3)} - \frac{U_{cc}A_{cc}}{M_{cc}C_v}\right)\begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}^\top z^{(3)} + \frac{\rho_a}{M_{cc}C_v}\left((h_{fg} - C_{pa})(w_z - w_{ao}) \; -C_{pa}x_1^{(3)}\right)\begin{bmatrix} 1 \\ 1 \end{bmatrix}^\top u_z^{(3)} \quad (68)$$

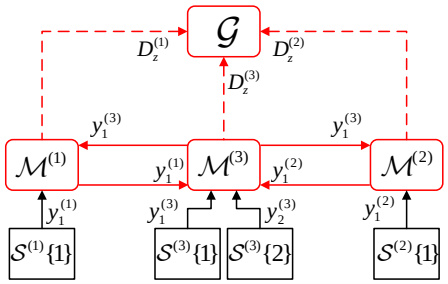


Fig. 4. Distributed sensor fault diagnosis architecture.

$\mathcal{F}_{c_4}^z = \left\{f_1^{(1)}, f_1^{(2)}\right\}$, $\mathcal{F}_{c_5}^z = \left\{f_1^{(1)}, f_1^{(3)}\right\}$, $\mathcal{F}_{c_6}^z = \left\{f_1^{(2)}, f_1^{(3)}\right\}$ and $\mathcal{F}_{c_7}^z = \left\{f_1^{(1)}, f_1^{(2)}, f_1^{(3)}\right\}$.

	$f_1^{(I)}$	$f_1^{(3)}$	$\{f_1^{(I)}, f_1^{(3)}\}$
$\mathcal{E}^{(I)}$	1	1	1

TABLE I
SENSOR FAULT SIGNATURE MATRIX $F^{(I)}$ OF $\mathcal{M}^{(I)}$, $I = 1, 2$.

	$\mathcal{F}_{c_1}^{(3)}$	$\mathcal{F}_{c_2}^{(3)}$	$\mathcal{F}_{c_3}^{(3)}$	$\mathcal{F}_{c_4}^{(3)}$	$\mathcal{F}_{c_5}^{(3)}$	$\mathcal{F}_{c_6}^{(3)}$	$\mathcal{F}_{c_7}^{(3)}$
$\mathcal{E}^{(3,1)}$	1	0	1	1	1	1	1
$\mathcal{E}^{(3,2)}$	0	1	1	0	0	1	1

TABLE II
SENSOR FAULT SIGNATURE MATRIX $F^{(3)}$ OF $\mathcal{M}^{(3)}$.

	$\mathcal{F}_{c_1}^z$	$\mathcal{F}_{c_2}^z$	$\mathcal{F}_{c_3}^z$	$\mathcal{F}_{c_4}^z$	$\mathcal{F}_{c_5}^z$	$\mathcal{F}_{c_6}^z$	$\mathcal{F}_{c_7}^z$
$\mathcal{E}^{(1)}$	1	0	*	1	1	*	1
$\mathcal{E}^{(2)}$	0	1	*	1	*	1	1
$\mathcal{E}^{(3)}$	*	*	1	*	1	1	1

TABLE III
SENSOR FAULT SIGNATURE MATRIX F^z IN $\mathcal{G}$.

In this example, the modeling uncertainty of $\eta^{(I)}(t)$ for $I = 1, 2$, is simulated as $\eta^{(I)}(t)$=5%$Y_1^{(I)}\sin(2\pi\nu t)$ and the random, uniformly bounded noise of each sensor $\mathcal{S}^{(I)}\{j\}$, characterized by (3), is simulated as $\bar{d}_j^{(I)} = 3\%Y_j^{(I)}$, $I = 1, 2$, $j = 1, 2$, where $Y_j^{(I)}$, is the steady state value of $y_j^{(I)}$ under healthy conditions ($Y_1^{(I)} = 24$, $I = 1, 2$, $Y_1^{(3)} = 10$, $Y_2^{(3)} = 5$). Two fault scenarios were simulated; in the first scenario, $f_1^{(1)}$ and $f_1^{(2)}$ occur simultaneously at $T_{f_1}^{(1)} = T_{f_1}^{(2)} = 5$ 000sec; in the second scenario, $f_1^{(3)}$ and $f_2^{(3)}$ occur simultaneously at $T_{f_1}^{(3)} = T_{f_2}^{(3)} = 5$ 000sec. In both scenarios, abrupt, bias sensor faults were simulated, where $\phi_1^{(1)}(t) = 15\%Y_1^{(1)}$, $\phi_1^{(2)}(t) = 15\%Y_1^{(2)}$, $\phi_1^{(3)}(t) = 10\%Y_1^{(3)}$ and $\phi_2^{(3)}(t) = 10\%Y_2^{(3)}$.

Figure 5 and 6 illustrate the result of the distributed SFDI technique for the first and second simulated scenario, respectively. According to Fig. 5, at the time instant 5000 sec, the diagnosis set generated by the agents $\mathcal{M}^{(1)}$ and $\mathcal{M}^{(2)}$ is $\mathcal{D}_s^{(1)}(t) = \left\{\{f_1^{(1)}\}, \{f_1^{(3)}\}, \{f_1^{(1)}, f_1^{(3)}\}\right\}$ and $\mathcal{D}_s^{(2)}(t) = \left\{\{f_1^{(2)}\}, \{f_1^{(3)}\}, \{f_1^{(2)}, f_1^{(3)}\}\right\}$ for $t \geq 5000$sec, while the

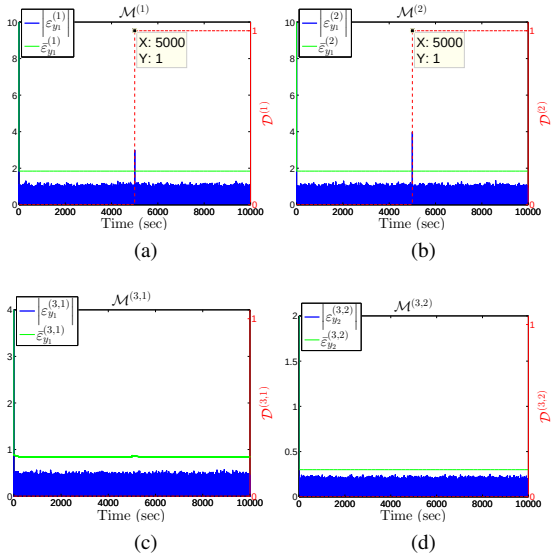


Fig. 5. Distributed SFDI for the first sensor fault scenario. The observed patterns $D^{(1)}(t)$, $D^{(2)}(t)$, and $D^{(3)} = [D^{(3,1)}(t), D^{(3,2)}(t)]^\top$ are respectively compared to the matrices $F^{(1)}$, $F^{(2)}$ and $F^{(3)}$ (Tables I and II).

agent $\mathcal{M}^{(3)}$ does not detect any fault. The agents $\mathcal{M}^{(1)}$ and $\mathcal{M}^{(2)}$ cannot isolate sensor faults, since, as shown in Table I, a violation of the ARR may be due to any sensor fault combination. Hence, their decisions along with the decision of agent $\mathcal{M}^{(3)}$ are combined and processed by applying the global decision logic. The observed pattern of propagated sensor faults is $D_z(t) = [1, 1, 0]^\top$ is consistent with the pattern F_4^z of the matrix F^z displayed in Table III, implying that the diagnosis set is $\mathcal{D}_s^z = \left\{\{f_1^{(1)}, f_1^{(2)}\}\right\}$. Combining the diagnosis sets $\mathcal{D}_s^{(1)}(t)$, $\mathcal{D}_s^{(2)}(t)$ and $\mathcal{D}_s^z(t)$, the agents $\mathcal{M}^{(1)}$ and $\mathcal{M}^{(2)}$ infer that local sensor faults ($f_1^{(1)}$ and $f_1^{(2)}$) have occurred, excluding the occurrence of $f_1^{(3)}$.

According to Fig. 6, at the time instant 5000 sec $D^{(3)} = [1, 1]^\top$, thus based on Table II, the diagnosis set is $\mathcal{D}_s^{(3)}(t) = \left\{\{f_1^{(3)}, f_2^{(3)}\}, \{f_z^{(3)}, f_2^{(3)}\}, \{f_z^{(3)}, f_1^{(3)}, f_2^{(3)}\}\right\}$, while the agents $\mathcal{M}^{(1)}$ and $\mathcal{M}^{(2)}$ do not detect any fault. Given $\mathcal{D}_s^{(3)}(t)$, the agent $\mathcal{M}^{(3)}$ isolates the sensor fault $f_2^{(3)}$, but it cannot infer the occurrence of the local fault $f_1^{(3)}$ and/or propagated faults $f_1^{(1)}, f_1^{(2)}$. Based on the decisions of the three monitoring agents, the observed pattern of propagated sensor faults is $D_z(t) = [0, 0, 1]^\top$ for $t \geq 5000$, which is compared to the columns of the matrix F^z displayed in Table III. The observed pattern $D_z(t)$ is consistent with the theoretical pattern F_3^z, implying that the diagnosis set is $\mathcal{D}_s^z = \left\{f_1^{(3)}\right\}$. Combining the diagnosis set $\mathcal{D}_s^3$ and $\mathcal{D}_s^z$, we infer that sensor faults $f_1^{(3)}$ and $f_2^{(3)}$ have occurred. It is noted that the effects of sensor fault $f_1^{(3)}$ were too low to be detectable from the agents $\mathcal{M}^{(1)}$ and $\mathcal{M}^{(2)}$, as presented in Fig. 6a and 6b (correspondingly for $f_1^{(1)}$ and $f_1^{(2)}$ in the first simulation scenario, as shown in Fig. 5c), since the agents $\mathcal{M}^{(1)}$ and $\mathcal{M}^{(2)}$ do not detect any sensor fault (correspondingly for $\mathcal{M}^{(3)}$ in the first simulation scenario). This happens due to the type of interconnection function $h^{(I)}$ that weakens the effects of

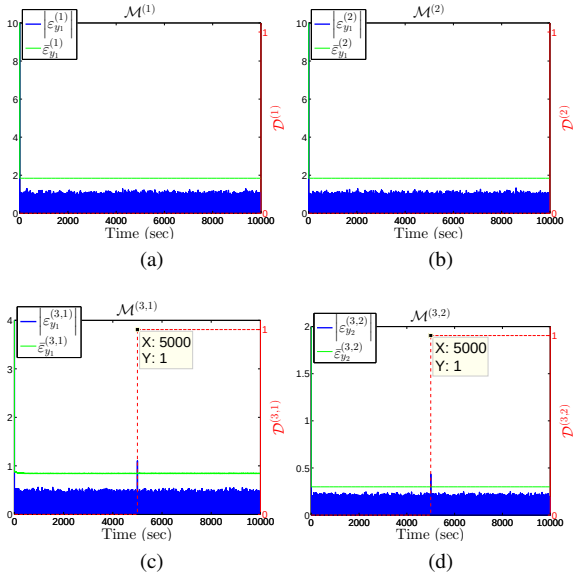


Fig. 6. Distributed SFDI for the second sensor fault scenario. The observed patterns $D^{(1)}(t)$, $D^{(2)}(t)$, and $D^{(3)} = [D^{(3,1)}(t), D^{(3,2)}(t)]^\top$ are respectively compared to the matrices $F^{(1)}$, $F^{(2)}$ and $F^{(3)}$ (Tables I and II).

fault $f_1^{(3)}$ on the agents $\mathcal{M}^{(1)}$ and $\mathcal{M}^{(2)}$, as described in Section VI (correspondingly for $h^{(3)}$ in the first scenario).

VIII. Conclusion

In this paper, we presented a distributed sensor fault detection and isolation technique for interonnected, cyber-physical systems (CPS). The backbone of the proposed method is the pursuit of two-level diagnosis in the cyber superstratum. The first-level diagnosis is conducted by a bank of monitoring agents, while a global decision logic is applied for conducting the second-level diagnosis. Each monitoring agent is designed to diagnose multiple faults in the sensors of the corresponding, interconnected subsystem, while it is allowed to exchange information with its neighboring agents. The goal of the global decision logic is to isolate multiple sensor faults propagating through the information exchanged between CPS. The proposed methodology is analyzed with respect to propagated and local sensor fault effects on the decisions of the monitoring agents and the distributed sensor fault detectability. Future work will involve the application of the proposed scheme to large-scale examples of interconnected CPS, such as mobile robotics, intelligent transportation, smart buildings.

References

[1] P. J. Antsaklis, B. Goodwine, V. Gupta, M. J. McCourt, Y. Wang, P. Wu, M. Xia, H. Yu, and F. Zhu, "Control of cyberphysical systems using passivity and dissipativity based methods," *European Journal of Control*, vol. 19, no. 5, pp. 379 – 388, 2013.

[2] M. Blanke, M. Kinnaert, J. Lunze, and M. Staroswiecki, *Diagnosis and fault-tolerant control*. Springer Verlag, 2003.

[3] R. Isermann, *Fault-diagnosis systems: An introduction from fault detection to fault tolerance*. Springer Verlag, 2006.

[4] X. Yan and C. Edwards, "Robust decentralized actuator fault detection and estimation for large-scale systems using a sliding mode observer," *International Journal of Control*, vol. 81, no. 4, pp. 591–606, 2008.

[5] X. Zhang and Q. Zhang, "Distributed fault diagnosis in a class of interconnected nonlinear uncertain systems," *International Journal of Control*, vol. 85, no. 11, pp. 1644–1662, 2012.

[6] S. Klinkhieo, R. J. Patton, and C. Kambhampati, "Robust FDI for FTC coordination in a distributed network system," in *16th IFAC World Congress*, Seoul, Korea, 2008, pp. 13 551–13 556.

[7] R. M. G. Ferrari, T. Parisini, and M. M. Polycarpou, "Distributed fault detection and isolation of large-scale discrete-time nonlinear systems: An adaptive approximation approach," *IEEE Transactions on Automatic Control*, vol. 57, no. 2, pp. 275–290, 2012.

[8] 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.

[9] H. Ferdowsi, D. L. Raja, and S. Jagannathan, "A decentralized fault detection and prediction scheme for nonlinear interconnected continuous-time systems," in *The 2012 International Joint Conference on Neural Networks (IJCNN)*, 2012, pp. 1–7.

[10] S. Indra, E. Chanthery *et al.*, "Decentralized diagnosis with isolation on request for spacecraft," in *8th IFAC SAFEPROCESS*, Mexico City, Mexico, 2012, pp. 283–288.

[11] 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.

[12] M. Daigle, X. Koutsoukos, and G. Biswas, "Distributed diagnosis in formations of mobile robots," *IEEE Transactions on Robotics*, vol. 23, no. 2, pp. 353–369, 2007.

[13] M. Davoodi, K. Khorasani, H. Talebi, and H. Momeni, "Distributed fault detection and isolation filter design for a network of heterogeneous multiagent systems," *IEEE Transactions on Control Systems Technology*, vol. 22, no. 3, pp. 1061–1069, 2014.

[14] Q. Zhang and X. Zhang, "Distributed sensor fault diagnosis in a class of interconnected nonlinear uncertain systems," in *8th IFAC SAFEPROCESS*, Mexico City, Mexico, 2012, pp. 1101–1106.

[15] ——, "Distributed sensor fault diagnosis in a class of interconnected nonlinear uncertain systems," *Annual Reviews in Control*, vol. 37, pp. 170 – 179, 2013.

[16] V. Reppa, M. M. Polycarpou, and C. G. Panayiotou, "Multiple sensor fault detection and isolation for large-scale interconnected nonlinear systems," in *European Control Conference (ECC2013)*, Zurich, Switzerland, 2013, pp. 1952–1957.

[17] ——, "Decentralized isolation of multiple sensor faults in large-scale interconnected nonlinear systems," *IEEE Transactions on Automatic Control*, conditionally accepted.

[18] ——, "A distributed detection and isolation scheme for multiple sensor faults in interconnected nonlinear systems," in *52nd Conference on Decision and Control (CDC2013)*, Florence, Italy, 2013, pp. 4991–4996.

[19] Q. Zhang, X. Zhang, M. M. Polycarpou, and T. Parisini, "Distributed sensor fault detection and isolation for multimachine power systems," *International Journal of Robust and Nonlinear Control*, vol. 24, no. 8-9, pp. 1403–1430, 2014.

[20] F. Zhu and Z. Han, "A note on observers for lipschitz nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 47, no. 10, pp. 1751–1754, 2002.

[21] R. Rajamani, "Observers for lipschitz nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 43, pp. 397–401, 1998.

[22] P. A. Ioannou and J. Sun, *Robust Adaptive Control*. Prentice-Hall, 1995.

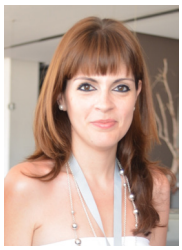
[23] V. Reppa, M. Polycarpou, and C. Panayiotou, "Adaptive approximation for multiple sensor fault detection and isolation of nonlinear uncertain systems," *IEEE Transactions on Neural Networks and Learning Systems,*, vol. 25, no. 1, pp. 137–153, 2014.

[24] M. Cordier, P. Dague, F. Lévy, J. Montmain, M. Staroswiecki, and L. Travé-Massuyès, "Conflicts versus analytical redundancy relations: a comparative analysis of the model based diagnosis approach from the artificial intelligence and automatic control perspectives," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 34, no. 5, pp. 2163–2177, 2004.

[25] V. Puig, A. Stancu, and J. Quevedo, "Robust fault isolation using nonlinear interval observers: the DAMADICS benchmark case study," in *16th IFAC World Congress*, 2005, pp. 1850–1855.

[26] J. Koscielny, M. Bartys, and M. Syfert, "Method of multiple fault isolation in large scale systems," *IEEE Transactions on Control Systems Technology*, vol. 20, no. 5, pp. 1302–1310, 2012.

[27] V. Reppa, P. Papadopoulos, M. M. Polycarpou, and C. G. Panayiotou, "Distributed detection and isolation of sensor faults in HVAC systems," in *21st Mediterranean Conference on Control and Automation (MED)*, 2013, pp. 401–406.



Vasso Reppa received the Diploma and Ph.D. degree in electrical and computer from the University of Patras, Patras, Greece, in 2004 and 2010, respectively. In parallel, she was an External Scientific Collaborator with Patras Scientific Park S.A., Patras, Greece, from 2006 to 2008. She was a Student Intern in the Storage Technologies Department, IBM Zurich Research Laboratory, Rüschlikon, Switzerland, in 2009. From 2011 to 2013, she was a Post-Doctoral Researcher with the KIOS Research Center for Intelligent Systems and Networks, University of Cyprus, Nicosia, Cyprus. Her research interests include fault diagnosis, fault tolerant control, adaptive learning, set-membership identification, with applications to microelectromechanical systems and large-scale engineering systems. She was a Researcher in various Hellenic and European research and operational programmes.

In 2014, Dr. Reppa was awarded the Marie Curie Intra European Fellowship and is working as a post-doctoral researcher with Automatic Control Department, in Supélec, Gif-sur-Yvette, France.

Marios M. Polycarpou is a Professor of Electrical and Computer Engineering and the Director of the KIOS Research Center for Intelligent Systems and Networks at the University of Cyprus. His research and teaching interests are in intelligent systems and control, fault diagnosis, adaptive and cooperative control systems, computational intelligence, and distributed agents. He has published more than 250 articles in refereed journals, edited books and refereed conference proceedings, and co-authored 6 books. He is also the holder of 6 patents.

Prof. Polycarpou is a Fellow of the IEEE and served as the President of the IEEE Computational Intelligence Society (01/2012–12/2013). He has served as the Editor-in-Chief of the IEEE Transactions on Neural Networks and Learning Systems between 2004-2010. He participated in 60 research projects/grants, funded by several agencies and industry in Europe and the United States. In 2011, Prof. Polycarpou was awarded the prestigious European Research Council (ERC) Advanced Grant.

Christos G. Panayiotou has received a B.Sc. and a Ph.D. degree in Electrical and Computer Engineering from the University of Massachusetts, Amherst, in 1994 and 1999 respectively. From 1999 to 2002 he was a Research Associate at the Center for Information and System Engineering (CISE) and the Manufacturing Engineering Department at Boston University. Since 2002-2003 he has been with the Department of Electrical and Computer Engineering, University of Cyprus where he is currently an Associate Professor. He is also a founding member of the KIOS Research Center on Intelligent Systems and Networks. His research interests include wireless, ad hoc and sensor networks, distributed control systems, fault diagnosis and fault tolerant systems, computer communication networks, optimization and control of discrete-event systems, resource allocation, simulation. Christos is an Associate Editor for the Conference Editorial Board of the IEEE Control Systems Society, the Journal of Discrete-Event Dynamical Systems, and the European Journal of Control.