Scalable distributed sensor fault diagnosis for smart buildings

Papadopoulos, Panayiotis M.; Reppa, Vasso; Polycarpou, Marios M.; Panayiotou, Christos G.

DOI
10.1109/JAS.2020.1003123

Publication date
2020

Document Version
Final published version

Published in
IEEE/CAA Journal of Automatica Sinica

Scalable Distributed Sensor Fault Diagnosis for Smart Buildings

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

***Abstract*— The enormous energy use of the building sector and the requirements for indoor living quality that aim to improve occupants' productivity and health, prioritize Smart Buildings as an emerging technology. The Heating, Ventilation and Air-Conditioning (HVAC) system is considered one of the most critical and essential parts in buildings since it consumes the largest amount of energy and is responsible for humans comfort. Due to the intermittent operation of HVAC systems, faults are more likely to occur, possibly increasing eventually building's energy consumption and/or downgrading indoor living quality. The complexity and large scale nature of HVAC systems complicate the diagnosis of faults in a centralized framework. This paper presents a distributed intelligent fault diagnosis algorithm for detecting and isolating multiple sensor faults in large-scale HVAC systems. Modeling the HVAC system as a network of interconnected subsystems allows the design of a set of distributed sensor fault diagnosis agents capable of isolating multiple sensor faults by applying a combinatorial decision logic and diagnostic reasoning. The performance of the proposed method is investigated with respect to robustness, fault detectability and scalability. Simulations are used to illustrate the effectiveness of the proposed method in the presence of multiple sensor faults applied to a 83-zone HVAC system and to evaluate the sensitivity of the method with respect to sensor noise variance.**

***Index Terms*—Building automation, fault diagnosis, fault location, smart homes.**

I. Introduction

A. Motivation

ACCORDING to the National Human Activity Pattern Survey (NHAPS), an average person in USA spends 86.9% of his/her life indoors [1]. The motivation for the development of smart buildings is the need to increase the energy efficiency of buildings [2], and the reliability of building's automation process [3], while decreasing the risk of safety-critical conditions [4], [5]. Since humans spend so much of their time indoors, health and living quality highly depends on the indoor conditions related to humidity, temperature, quality of air and many more. These factors are closely related to the safe and reliable operation of the Heating, Ventilation and Air-Conditioning (HVAC) system.

HVAC systems are complex machines that consist of a large number of interconnected components. Faults in electrical and mechanical equipment such as sensors, wires, fans, valves, pumps of the HVAC system are inevitable due to its continual operation. Faults can increase the energy consumption and create discomfort conditions for occupants. The feedback control performance of the HVAC system depends on the availability and reliability of sensor measurements. Advances in wireless network technology and Internet of Things (IoT) technology have enhanced the availability of data in buildings [6]. However, data unreliability due to potential sensor faults can be a major drawback for the performance of the feedback control process.

B. Literature Survey

Fault diagnosis is a well-established procedure to discover anomalies in systems [7]. Currently, the industry of HVAC control systems uses *ruled-based* algorithms to diagnose anomalies during the operation of HVAC systems. The rules are formed by comparing sensor data or relations of sensor data with predefined constant thresholds obtained by experts; sometimes these are also called expert systems. Some examples of ruled-based fault diagnosis schemes for HVAC systems are: 1) the performance assessment rules that identify the mode of operation using specific relationships of measured information [8], [9]; and 2) the cause-effect graphs where the various operation modes of the system (both healthy and faulty modes) are represented as discrete events [10], [11]. The main weaknesses of rule-based fault diagnosis methods are that they are very specific to the system, can fail beyond the boundaries of the expertise incorporated in them, and are difficult to update [12]; in other words, ruled-based fault diagnosis methods do not employ any adaptability when the algorithm is applied to HVAC systems and buildings with different properties and/or parameters.

Intelligent fault diagnosis algorithms can be divided into two categories; *data-driven/data-mining* and *model-based* fault diagnosis algorithms. The former category includes traditional computational intelligence algorithms that originate from

Manuscript received December 10, 2019; revised January 17, 2020; accepted February 19, 2020. This work was supported by the European Union's Horizon 2020 Research and Innovation Programme (739551) (KIOS CoE). Recommended by Associate Editor Chengdong Li. *(Corresponding author: Panayiotis M. Papadopoulos.)*

Citation: P. M. Papadopoulos, V. Reppa, M. M. Polycarpou, and C. G. Panayiotou, "Scalable distributed sensor fault diagnosis for smart buildings," *IEEE/CAA J. Autom. Sinica*, vol. 7, no. 3, pp. 638–655, May 2020.

P. M. Papadopoulos, M. M. Polycarpou, and C. G. Panayiotou are with the KIOS Research and Innovation Center of Excellence, Department of Electrical and Computer Engineering, University of Cyprus, Nicosia 1678, Cyprus (e-mail: papadopoulos.panagiotis@ucy.ac.cy; mpolycar@ucy.ac.cy; christosp@ucy.ac.cy).

V. Reppa is with the Department of Maritime and Transport Technology, Delft University of Technology, Delft, 2628 CD, The Netherlands (e-mail: v.reppa@tudelft.nl).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JAS.2020.1003123

machine learning and pattern recognition. Most of the data-driven methods require historical data (i.e., database of sensor data) for training the fault decision rules. Amongst the popular data-driven methods are: principal component analysis (PCA) [13], [14], support vector machines (SVM) [15]–[18], neural networks (NN) [19], [20], genetic algorithms (GA) [21], [22], fuzzy logic models [23], [24], etc.

Model-based fault diagnosis methods can be classified according to the type of model, that is; *statistical* and *state-space* models. Statistical models use data to identify a simple model such as: autoregressive model with exogenous inputs (ARX) [25], autoregressive moving average model with exogenous inputs (ARMAX) [26], [27], fast Fourier transform (FFT) model [28]. The statistical models try to predict the output of the system during operation using techniques such as average error of residuals. Statistical analysis employs simplistic models that require a training interval to obtain the corresponding model parameters and the state of the system (e.g., temperature). The state is represented as a random variable that is a linear combination of its previous values. In order to obtain a valid prediction of system's state using statistical models, an adequate training and knowledge of the initial state of the system are required. The latter subcategory corresponds to the state estimation models that perform online learning of the state based on the real-time data of the system. Some examples of fault diagnosis algorithms based on state estimation models are: the Kalman filtering approach [29] and the observer-based estimation schemes [30], [31]. State estimation techniques allow the utilization of nonlinear representations of the system dynamics that provide a more realistic characterization of the heat transfer processes, compared to the aforementioned methods that employ an approximated model.

The utilization of analytical models describing the physical environment of the building and the HVAC system is challenging because of: 1) the possibly unknown thermal properties; 2) the large number of physically interconnected building zones; and 3) the complexity of the electromechanical part of the HVAC system. A recently established European legislative framework about energy performance of buildings directive, includes the issuing of buildings energy performance certificates. Thermal quality of the building envelope (e.g., structure, material values) and energy efficiency of building equipment (such as heating and cooling systems) are required for the certificate, thus building's and HVAC system's thermal properties can be available for the state-space representation of system's dynamics.

Model-based fault diagnosis algorithms can be applied without necessitating any training period compared to data-driven methods that cannot guarantee the robustness of the decision outcome (i.e., detection, isolation), since the decision highly depends on the training set, that may no cover all the possible sensor fault scenarios. Data-driven methods commonly use a fixed, pre-designed detection threshold calculated using a training set [14], [32]–[34]. Hence, false alarms may be triggered in the presence of an event that was not included in the training set. Further, with respect to the fault isolation procedure, data-driven methods necessitate historical data of faulty situations (that are typically difficult to obtain) in order to build the isolation logic.

Most works in the literature of model-based fault diagnosis address the problem of fault diagnosis for single-zone HVAC systems [22], [35]–[37]. Only few of them deal with fault diagnosis in multi-zone HVAC systems, instead assuming that there is no heat transfer between zones [38]–[41], or assuming that there is heat transfer between zones only through walls (no doors) [40], [42]. Albeit simpler, these models may be unrealistic in practice because they cannot capture the variant heat transfer between building zones due to the presence of doors, which is a stronger physical interconnection than the heat transfer through walls. Ignoring the strong physical interconnections between building zones can result in high modeling errors, which may cause false alarms or missed fault detection.

In large-scale buildings, the utilization of a global model describing the entire building system can be prohibitive for the design of a model-based fault diagnosis technique. Exploiting the distributed topology of the building system, every fault diagnosis agent can be designed to monitor a single building zone and to execute the fault isolation process locally, while taking into account faults that affect part of the building system and not the entire system [43]. This strategy is effective for handling the problem of the occurrence of multiple homogeneous or heterogeneous faults [31], [41]. Moreover, the distributed architecture can be scalable in the case that the building topology evolves, since a new fault diagnosis agent dedicated to the new building part can be augmented, similar to a plug-and-play approach [44]. With the spatially distributed deployment of the fault diagnosis agents, there is no central point for executing the fault diagnosis process that corresponds to a "single point of failure." This is especially important in safety-critical buildings such as hospitals, schools, and other public buildings.

In previous works, the authors designed and evaluated a distributed approach for sensor fault detection and isolation [39], [40] and distributed sensor fault accommodation [42], [45] in HVAC systems. More recently, the authors proposed a distributed methodology to identify and isolate actuator and sensor faults [31]. However, the performance of the fault diagnosis methods in the aforesaid studies was evaluated in small-scale buildings. Moreover, the analysis of the detectability was only examined in a HVAC system with separated zones where the heat transfer between zones (through walls and/or doors) was not considered in system's dynamics [41]. Modeling the heat transfer between zones leads to non-linear, non-Lispschitz dynamic terms that can create a more realistic model and thus less conservative fault detection thresholds. This can improve the detectability aspect of the algorithm since the modelling error is reduced and moreover can avert any false alarms caused by the event of an opened door. However, dealing with hard nonlinearities creates challenges with the design and analysis.

C. Contribution

The goal and main contribution of this work is the design of a scalable distributed model-based method for diagnosing

multiple sensor faults in large-scale HVAC systems, while taking into account interconnected building zones through walls and doors. Based on the topology of the HVAC system and the building zones, the overall system is divided into interconnected subsystems. A sensor set $\mathcal{S}^{(i)}$ collects the measurements of each subsystem. A local sensor fault diagnosis agent $\mathcal{M}^{(i)}$ is designed to monitor the corresponding sensor set $\mathcal{S}^{(i)}$ and to detect and isolate single and multiple sensor faults based on local state estimation obtained using local information and information transmitted from its neighboring agents (e.g., control inputs, sensor measurements) as illustrated in Fig. 1. Each dedicated sensor fault diagnosis agent $\mathcal{M}^{(i)}$ is comprised of a distributed sensor fault detection module and a distributed sensor fault isolation module. The former is responsible for detecting the occurrence of sensor faults in the monitored subsystem and/or its neighboring interconnected subsystems. A local detection signal is generated by comparing the residual (that corresponds to the discrepancies between the output and the expected output) with the corresponding adaptive threshold (designed to bound the residual under healthy conditions).

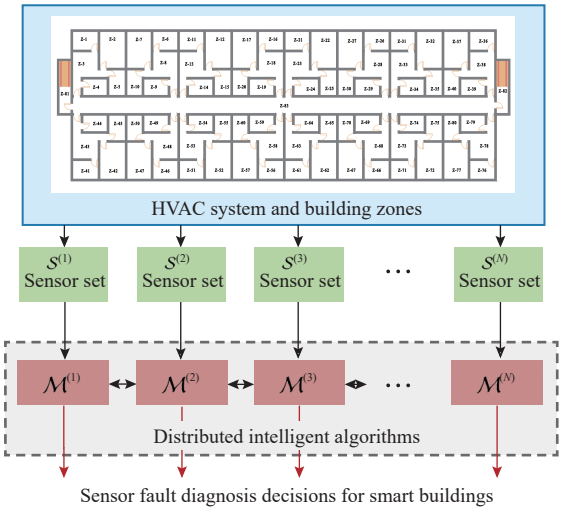


Fig. 1. Architecture of the distributed sensor fault diagnosis scheme for smart buildings.

Based on the local state estimation, each agent can detect sensor faults affecting either the local or the neighbouring subsystems. The distributed sensor fault isolation module, which is activated based on the local detection decision, takes into consideration the connectivity (due to the exchange of information between the distributed diagnosis agents) in order to construct a fault signature matrix that can eliminate a number of possible sensor faults and under some conditions can pinpoint the exact location of sensor faults. The performance analysis of the proposed method is provided with respect to robustness, fault detectability and scalability, taking into account modelling uncertainties and strong physical interconnections between the building zones that can improve the fault detectability of the algorithm since the modelling error is reduced and moreover can avert false detection alarms caused by the event of an opened door. Finally, simulation results generated by the application of the proposed method to a large-scale HVAC building system show its effectiveness.

D. Paper Organization

The paper is organized as follows. Section II presents the problem formulation that consists of the description of the HVAC system and the network configuration. In Section III the architecture of the distributed sensor fault diagnosis methodology is given. The performance of the method is analyzed in Section IV. Section V presents some simulation results of a multiple sensor fault scenario for 83-zone HVAC system, followed by some concluding remarks in Section VI.

E. Nomenclature

$T_{st}(t)$ (°C) water temperature in the storage tank
$T_{z_i}(t)$ (°C) i-th zone air temperature
N number of zones
$T_{pl}(t)$ (°C) plenum (duct) temperature
$T_o(t)$ (°C) source heat temperature of the heat pump
$P_s(T_{st}(t))$ performance coefficient of the heat pump
$\widetilde{T}_{st}(t)$ (°C) disturbances affecting the water temperature dynamics due to, e.g., defective thermal insulation of the storage tank
$P_{\max}$ rated maximum value of $P_s(T_{st}(t))$
$\Delta T_{\max}$ (°C) maximum temperature difference for the heat pump
C_{st} (kJ/°C) heat capacity of the storage tank
$U_{i,\max}$ (kg/h) maximum mass flow rate of hot water through the coil placed at the i-th zone
$U_{st,\max}$ (kJ/h) heat pump rated capacity
a_{sz} (kJ/kg °C) effectiveness of the heating coil
a_{st} (kJ/kg °C) heat loss coefficient of storage tank from exterior surfaces
$u_i(t)$ mass flow rate of hot water flowing in the coil of i-th zone
$u_{st}(t)$ normalized energy in the heat pump
$T_{i1}(t)$ (°C) known temperature of the surface node of the mass wall in the i-th zone
$T_{amb}(t)$ (°C) known ambient temperature
$\widetilde{T}_{z_i}(t)$ (°C) temperature dynamics of the i-th zone due to presence of appliances, occupants, lights
h (W/m^2 °C) heat transfer coefficient due to the presence of walls
A_{w_i} (m^2) surface area of the mass wall
C_{z_i} (kJ/°C) air heat capacity of the i-th zone
a_{z_i} (kJ/h °C) heat loss coefficient of the i-th zone
$a_{z_{i,j}}$ (kJ/h °C) inter-zone heat loss coefficient between i-th and j-th zone due to the presence of walls
$A_{d_{i,j}}$ (m^2) area of the door connecting i-th and j-th zone
Σ^s, $\Sigma^{(i)}$ subsystems of storage tank and i-th zone
A^s, $A^{(i)}$ linearized parts of subsystems Σ^s, $\Sigma^{(i)}$
$g^s(\cdot)$, $g^{(i)}(\cdot)$ bilinear terms of subsystems Σ^s, $\Sigma^{(i)}$
$h^s(\cdot)$, $h^{(i)}(\cdot)$ interconnection terms of subsystems Σ^s, $\Sigma^{(i)}$
η^s, $\eta^{(i)}$ known exogenous inputs of subsystems Σ^s, $\Sigma^{(i)}$
$r^s(\cdot)$, $r^{(i)}(\cdot)$ uncertainty terms of subsystems Σ^s, $\Sigma^{(i)}$
$\mathcal{K}_i$ set of indices of zones that are interconnected with the i-th zone

$y^s_{\text{ref}}, y^{(i)}_{\text{ref}}$ reference signals for the states T_{st}, T_{z_i}
$n^s, n^{(i)}$ unknown measurement noise of sensors $\mathcal{S}^s, \mathcal{S}^{(i)}$
$f^s, f^{(i)}$ unknown sensor faults of sensors $\mathcal{S}^s, \mathcal{S}^{(i)}$
$y^s, y^{(i)}$ measurements of sensors $\mathcal{S}^s, \mathcal{S}^{(i)}$
$\hat{T}_{st}, \hat{T}_{z_i}$ estimation of states T_{st}, T_{z_i}
$\varepsilon^s, \varepsilon^{(i)}$ residuals of fault diagnosis agents $\mathcal{M}^s, \mathcal{M}^{(i)}$
$\bar{\varepsilon}^s, \bar{\varepsilon}^{(i)}$ adaptive thresholds of fault diagnosis agents $\mathcal{M}^s, \mathcal{M}^{(i)}$

II. Problem Formulation

This section first presents the dynamics of a multi-zone HVAC system and then characterizes the multi-zone HVAC model as a network of interconnected subsystems that will be used to design the distributed sensor fault diagnosis scheme.

A. HVAC System Description

This subsection presents the modeling of a multi-zone HVAC system, which is an extended version of the model presented in [46], [47] using terms from [48]. The overall modeling approach is illustrated in Fig. 2(a) using a simple example of a 5-zone HVAC system with heating operation. The electromechanical part of the system consists of a hot water unit, e.g., heat pump, condenser, storage tank (orange box in Fig. 2(a)). The hot water from the storage tank is circulated in the fan-coil units located in the plenum of each zone (black boxes in Fig. 2(a)) and then returns back to the storage tank. The same structure of the HVAC system can be used also for cooling operation by replacing the heat pump with a chiller. The water temperature in the storage tank is described by the thermal-mass balance equation expressed as

$$\begin{aligned}\frac{dT_{st}(t)}{dt} &= \frac{U_{st,\max}}{C_{st}} P_s(T_{st}(t)) u_{st}(t) - \frac{a_{st}}{C_{st}}(T_{st}(t) - T_{pl}(t)) \\ &\quad + \frac{a_{sz}}{C_{st}} \sum_{i\in\mathbb{N}} u_i(t) U_{i,\max}(T_{z_i}(t) - T_{st}(t)) + \frac{a_{st}}{C_{st}} \widetilde{T}_{st}(t) \end{aligned} \quad (1)$$

where

$$\begin{aligned} P_s(T_{st}(t)) &= \begin{cases} 1 + (P_{\max} - 1)\left(1 - \dfrac{\Delta T(t)}{\Delta T_{\max}}\right), & \Delta T(t) \le \Delta T_{\max} \\ 1, & \Delta T(t) > \Delta T_{\max} \end{cases} \\ \Delta T(t) &= T_{st}(t) - T_o(t) \end{aligned} \quad (2)$$

with $i \in \mathbb{N}$, $\mathbb{N} = \{1, \ldots, N\}$.

The i-th zone temperature dynamics can be described as

$$\begin{aligned}\frac{dT_{z_i}(t)}{dt} &= \frac{U_{i,\max} a_{sz}}{C_{z_i}}(T_{st}(t) - T_{z_i}(t)) u_i(t) \\ &\quad - \frac{a_{z_i}}{C_{z_i}}(T_{z_i}(t) - T_{amb}(t)) - \frac{hA_{w_i}}{C_{z_i}}(T_{i1}(t) - T_{z_i}(t)) \\ &\quad - \frac{1}{C_{z_i}} \sum_{j\in\mathcal{K}_i} a_{z_i,j}(T_{z_i}(t) - T_{z_j}(t)) + \frac{a_{z_i}}{C_{z_i}} \widetilde{T}_{z_i}(t) \\ &\quad + \frac{\rho_{\text{air}} C_p \sqrt{2(C_p - C_v)}}{C_{z_i}} \Big(\sum_{j\in\mathcal{K}_i} \text{sgn}(T_{z_j}(t) - T_{z_i}(t)) \\ &\quad \times A_{d_{i,j}} \max(T_{z_i}(t), T_{z_j}(t)) \sqrt{\left|T_{z_j}(t) - T_{z_i}(t)\right|} \Big) \end{aligned} \quad (3)$$

with $j \in \mathcal{K}_i$, $\mathcal{K}_i = \{j : a_{z_{ij}} \neq 0\}$. It is noted that $\mathcal{K}_i$ is the set that consists of the indices of zones that are interconnected with the i-th zone.

The objective is to develop a distributed estimation scheme using real-time sensor information for detecting and isolating abnormal behavior in the operation of the HVAC systems produced by the presence of sensor faults. The aforementioned thermal-mass balance equations are used in order to incorporate available modelling information in the estimation scheme (instead of using completely black-box methods). Also, it is important to note that the thermal-mass balance equations are not assumed to be completely accurate since the dynamics are characterised by unknown or uncertain terms.

B. Network Configuration

For design purposes, the thermal dynamics of the multi-zone HVAC system is characterized as a network of $N+1$ interconnected subsystems denoted by Σ^s, $\Sigma^{(1)}, \ldots, \Sigma^{(N)}$, where Σ^s represents the temperature dynamics of the storage tank, given in (1) and (2), and $\Sigma^{(i)}$, $i \in \mathbb{N}$, represents the temperature dynamics of the i-th building zone, decribed in (3). The subsystem Σ^s can be expressed as

$$\begin{aligned}\Sigma^s : \ \dot{T}_{st}(t) &= A^s T_{st}(t) + g^s(T_{st}(t)) u_{st}(t) + \eta^s(T_{pl}(t)) \\ &\quad + r^s(t) + h^s(T_{st}(t), T_z(t), u(t)) \\ &\triangleq \chi^s(T_{st}(t), T_z(t), u(t)) \end{aligned} \quad (4)$$

where $T_{st} \in \mathbb{R}$ represents the local state of subsystem Σ^s and $u_{st} \in \mathbb{R}$ denotes the local control input of subsystem Σ^s. The terms T_{pl} and T_o represent uncontrollable but known exogenous inputs. The variable $r^s(t) \triangleq \dfrac{a_{st}}{C_{st}} \widetilde{T}_{st}(t)$ models unknown inputs affecting the water temperature dynamics of the storage tank. The vector $T_z \triangleq \left[T_{z_1}, \ldots, T_{z_N}\right]$ is the interconnection vector that includes the states of neighboring subsystems (temperatures of all building zones), where T_{z_i} is the air temperature of the interconnected building zone i (i.e., state of subsystem $\Sigma^{(i)}$), and $u \triangleq [u_1, \ldots, u_N]$ is a vector that collects the control inputs of all building zones, where u_i is the control input of subsystem $\Sigma^{(i)}$. The terms g^s and h^s describe the nonlinear local and interconnection dynamics of subsystem Σ^s, respectively, and are defined as

$$g^s(T_{st}) = \frac{U_{st,\max}}{C_{st}} P_s(T_{st}) \quad (5)$$

$$h^s(T_{st}, T_z, u) = \frac{a_{sz}}{C_{st}} \sum_{i\in\{1,\ldots,N\}} U_{i,\max}(T_{st} - T_{z_i}) u_i. \quad (6)$$

The constant A^s is defined as $A^s = -a_{st}/C_{st}$ and $\eta^s(T_{pl}) = \dfrac{a_{st}}{C_{st}} T_{pl}$. The defined function χ^s collects all the dynamics of T_{st}. Each subsystem $\Sigma^{(i)}$ for all $i \in \mathbb{N}$, is interconnected with subsystems Σ^s and $\Sigma^{(j)}$, $j \in \mathcal{K}_i$, described by

$$\begin{aligned}\Sigma^{(i)} : \ \dot{T}_{z_i}(t) &= A^{(i)} T_{z_i}(t) + g^{(i)}(T_{st}(t), T_{z_i}(t)) u_i(t) + r^{(i)}(t) \\ &\quad + \eta^{(i)}(T_{i1}(t), T_{amb}(t)) + h^{(i)}(T_{z_i}(t), T_{\mathcal{K}_i}(t)) \\ &\triangleq \chi^{(i)}(T_{z_i}(t), T_{st}(t), T_{\mathcal{K}_i}(t), u_i(t)) \end{aligned} \quad (7)$$

where $T_{\mathcal{K}_i}(t) = [T_{z_j}(t) : j \in \mathcal{K}_i]^T$, $T_{\mathcal{K}_i}$ denotes a column vector

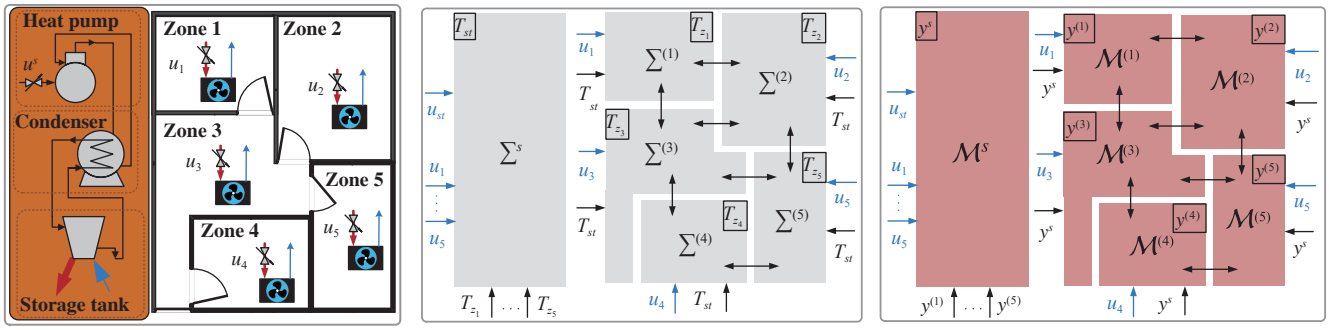


Fig. 2. The distributed sensor fault diagnosis architecture for a 5-zone HVAC system. From: (a) the physical system; (b) the mathematical model of the system; and (c) The distributed sensor fault diagnosis architecture. Specifically each subfigure shows: (a) Schematic representation of a multi-zone HVAC system that consists of the hot water unit (orange box) and the 5 building zones that are interconnected through walls and doors. The black rectangular boxes located in each zone represent the fan-coil units; (b) The subsystems network configuration of the 5-zone HVAC system. The black arrows denote the shared states between the interconnected subsystems; (c) the distributed sensor fault diagnosis agents $\mathcal{M}^s$ and $\mathcal{M}^{(1)},\ldots,\mathcal{M}^{(5)}$. The black arrows denote the exchange of information between the diagnosis agents.

of length $\text{card}(\mathcal{K}_i)$, where each element corresponds to the state T_{z_j} of the neighboring subsystem $\Sigma^{(j)}$, $j\in\mathcal{K}_i$. The variable $r^{(i)}(t)\triangleq\dfrac{a_{z_i}}{C_{z_i}}\widetilde{T}_{z_i}(t)$ models the unknown inputs of subsystem $\Sigma^{(i)}$, and $A^{(i)}=\dfrac{hA_{w_i}-a_{z_i}}{C_{z_i}}-\dfrac{1}{C_{z_i}}\sum_{j\in\mathcal{K}_i}a_{z_{i,j}}$. The terms $g^{(i)}$ and $h^{(i)}$ respectively denote the local and interconnection nonlinear dynamics of the subsystem $\Sigma^{(i)}$, i.e.,

$$g^{(i)}(T_{st},T_{z_i})=\sigma^{(i)}(T_{st}-T_{z_i}) \tag{8}$$

$$h^{(i)}(T_{z_i},T_{\mathcal{K}_i})=\frac{1}{C_{z_i}}\sum_{j\in\mathcal{K}_i}a_{z_i,j}T_{z_j}+p^{(i)}\Big(\sum_{j\in\mathcal{K}_i}\text{sgn}(T_{z_j}-T_{z_i}) \times A_{d_{i,j}}\max(T_{z_i},T_{z_j})\sqrt{\left|T_{z_j}-T_{z_i}\right|}\Big) \tag{9}$$

with $\sigma^{(i)}=\dfrac{U_{i,\max}a_{sz}}{C_{z_i}}$, $p^{(i)}=\dfrac{\rho_{\text{air}}C_p\sqrt{2(C_p-C_v)}}{C_{z_i}}$, and $\eta^{(i)}(T_{i1},T_{amb})=\dfrac{a_{z_i}}{C_{z_i}}T_{i1}-\dfrac{hA_{w_i}}{C_{z_i}}T_{amb}$. The function $\chi^{(i)}$ collects all the dynamics of T_{z_i}. The signals u_{st} in (4) and u_i in (7) are generated using a distributed feedback linearization control scheme based on some (differentiable) desired reference signals y^s_{ref} and $y^{(i)}_{\text{ref}}$ for the states T_{st} and T_{z_i}, respectively. An example of the network configuration of a 5-zone HVAC system is given in Fig. 2(b). The black arrows denote the shared states between the subsystems due to physical interconnections between the zones as well as between the storage tank and the zones.

The water temperature of Σ^s (storage tank) is measured by the sensor $\mathcal{S}^s$, characterized by

$$\mathcal{S}^s:\ y^s(t)=T_{st}(t)+n^s(t)+f^s(t) \tag{10}$$

where $y^s(t)$ is the sensor output and $n^s(t)$ is the unknown measurement noise. The output of the sensor $\mathcal{S}^{(i)}$ used to measure the air temperature of subsystem $\Sigma^{(i)}$ (zone i), i.e.,

$$\mathcal{S}^{(i)}:\ y^{(i)}(t)=T_{z_i}(t)+n^{(i)}(t)+f^{(i)}(t) \tag{11}$$

where $y^{(i)}(t)$ is the sensor output and $n^{(i)}(t)$ is the unknown measurement noise. The signals f^s and $f^{(i)}$ denote unknown sensor faults, defined as

$$f^s(t)=\beta^s(t-t^s_f)\phi^s(t-t^s_f) \tag{12}$$

$$f^{(i)}(t)=\beta^{(i)}(t-t^{(i)}_f)\phi^{(i)}(t-t^s_f) \tag{13}$$

where t^s_f, $t^{(i)}_f$ model the first time instants of fault occurrence, ϕ^s, $\phi^{(i)}$ are the fault functions and β^s, $\beta^{(i)}$ denote their time profiles where $\beta^s(t)\triangleq\beta_o(t;\alpha^s)$, $\beta^{(i)}(t)\triangleq\beta_o(t;\alpha^{(i)})$, with

$$\beta_o(t;\alpha)=\begin{cases}0, & t<0\\ 1-e^{-\alpha t}, & t\geq 0\end{cases} \tag{14}$$

where α^s, $\alpha^{(i)}$ being the time evolution rate of sensor faults f^s, $f^{(i)}$, respectively. Note that $\alpha\to\infty$ models an abrupt fault, while $\alpha\to 0$ describes a fault that evolves gradually.

The objective of this work is to design a scalable distributed methodology for detecting the faulty operation of the temperature sensors in the multi-zone HVAC system and isolating the location of the faulty sensors. Faults may occur at an unknown time in one or more building zones or in the electromechanical part of HVAC. The proposed methodology is designed taking into account the following assumption.

Assumption 1: For all $t\geq 0$, the modeling uncertainties $r^s(t)$, $r^{(i)}(t)$ and measurement noise $n^s(t)$, $n^{(i)}(t)$ are uniformly bounded such that $|r^s(t)|\leq\bar{r}^s$, $\left|r^{(i)}(t)\right|\leq\bar{r}^{(i)}$ and $|n^s(t)|\leq\bar{n}^s$, $\left|n^{(i)}(t)\right|\leq\bar{n}^{(i)}$, for all $i\in\mathbb{N}$.

Remark 1: The above assumption characterizes known bounds on the modeling uncertainty and measurement noise, which are required in order to distinguish between the occurrence of sensor faults and the presence of modeling uncertainties and measurement noise.

III. Distributed Sensor Fault Diagnosis Architecture

Based on the network of $N+1$ interconnected subsystems presented in Section II-B, a bank of distributed monitoring agents is developed. Fig. 2(c) illustrates the distributed structure of the sensor fault diagnosis agents (red boxes), dedicated to subsystem Σ^s (left) and to each subsystem $\Sigma^{(i)}$, $i\in\{1,\ldots,5\}$ (right). Every distributed sensor fault diagnosis agent is composed by the following two modules.

1) Sensor Fault Detection Module: Using the available (local and shared) sensor measurements and control inputs, an

estimator is designed based on the known nonlinear dynamics of its monitored subsystem. A residual, which corresponds to the deviation of the measured (observed) output of the monitored subsystem from the expected output, is generated on-line. Based on Assumption 1 and considering a healthy system, an adaptive threshold is designed to bound the residual at every time instant. Both the residual and the adaptive threshold are monitored on-line. The violation of the adaptive threshold indicates the presence of sensor faults and activates the sensor fault isolation module.

2) Sensor Fault Isolation Module: The local decision about the occurrence of sensor faults is processed in combination with the decisions of the neighboring agents, aiming at isolating multiple sensor faults.

A. Distributed Sensor Fault Detection Module

The design of the distributed sensor fault detection module includes the computation of residuals and adaptive thresholds, and the formulation of the sensor fault detection decision logic.

1) Residual Generation: The residual ε^s generated by the agent $\mathcal{M}^s$ is defined as

$$\varepsilon^s(t) = y^s(t) - \hat{T}_{st}(t) \tag{15}$$

where

$$\begin{aligned}\dot{\hat{T}}_{st}(t) = {} & (A^s - L^s)\hat{T}_{st}(t) + g^s(y^s(t))u_{st}(t) \\ & + h^s(y^s(t), y(t), u(t)) + \eta^s(T_{pl}(t)) + L^s y^s(t)\end{aligned} \tag{16}$$

where $\hat{T}_{st}$ is the estimation of the state T_{st} with $\hat{T}_{st}(0) = 0$, $y(t) = \left[y^{(1)}, \ldots, y^{(N)}\right]^T$, and L^s is the observer gain selected such that $(A^s - L^s)$ is negative. Let us define the estimation error as $\varepsilon_T^s(t) = T_{st}(t) - \hat{T}_{st}(t)$; then based on (4) and (16), ε_T^s satisfies

$$\begin{aligned}\dot{\varepsilon}_T^s(t) = {} & (A^s - L^s)\varepsilon_T^s(t) + \widetilde{g}^s(t)u_{st}(t) + \widetilde{h}^s(t) + r^s(t) \\ & - L^s\left(n^s(t) + f^s(t)\right)\end{aligned} \tag{17}$$

with $\widetilde{g}^s(t) \triangleq g^s(T_{st}(t)) - g^s(y^s(t))$, $\widetilde{h}^s(t) \triangleq h^s(T_{st}(t), T_z(t), u(t)) - h^s(y^s(t), y(t), u(t))$ and

$$\widetilde{g}^s = \frac{U_{st,\max}}{C_{st}}\left(P_s(T_{st}) - P_s(y^s)\right) \tag{18}$$

$$\widetilde{h}^s = \frac{a_{sz}}{C_{st}}\sum_{i\in\mathbb{N}} U_{i,\max}(n^{(i)} - n^s + f^{(i)} - f^s)u_i \tag{19}$$

where P_s is defined in (2). Based on (10), the residual ε^s defined in (15) can be re-written as

$$\varepsilon^s(t) = \varepsilon_T^s(t) + n^s(t) + f^s(t). \tag{20}$$

Remark 2: Given (10), (11), and (17)–(19), it yields that ε^s can be affected by a fault in sensor $\mathcal{S}^s$ and/or any sensor fault in $\mathcal{S}^{(i)}$, $i \in \mathbb{N}$.

The residual $\varepsilon^{(i)}$ generated by the monitoring agent $\mathcal{M}^{(i)}$, i.e.,

$$\varepsilon^{(i)}(t) = y^{(i)}(t) - \hat{T}_{z_i}(t) \tag{21}$$

where

$$\begin{aligned}\dot{\hat{T}}_{z_i}(t) = {} & \left(A^{(i)} - L^{(i)}\right)\hat{T}_{z_i}(t) + g^{(i)}(y^s(t), y^{(i)}(t))u_i(t) \\ & + \eta^{(i)}(T_{i1}(t), T_{amb}(t)) + h^{(i)}(y^{(i)}(t), y_{\mathcal{K}_i}(t)) + L^{(i)}y^{(i)}(t)\end{aligned} \tag{22}$$

where $\hat{T}_{z_i}(0) = 0$, $\hat{T}_{z_i}$ is the estimation of the state T_{z_i} (i-th zone air temperature), $y_{\mathcal{K}_i}(t) = \left[y^{(j)} : j \in \mathcal{K}_i\right]$, and $L^{(i)}$ is the observer gain selected such that $\left(A^{(i)} - L^{(i)}\right)$ is negative for all $i \in \mathbb{N}$. Let us define the estimation error as $\varepsilon_T^{(i)}(t) = T_{z_i}(t) - \hat{T}_{z_i}(t)$; then $\varepsilon_T^{(i)}$ satisfies

$$\begin{aligned}\dot{\varepsilon}_T^{(i)}(t) = {} & \left(A^{(i)} - L^{(i)}\right)\varepsilon_T^{(i)}(t) + \widetilde{g}^{(i)}(t)u_i(t) + \widetilde{h}^{(i)}(t) + r^{(i)}(t) \\ & - L^{(i)}\left(n^{(i)}(t) + f^{(i)}(t)\right)\end{aligned} \tag{23}$$

where $\widetilde{g}^{(i)}(t) \triangleq g^{(i)}(T_{st}(t), T_{z_i}(t)) - g^{(i)}(y^s(t), y^{(i)}(t))$, $\widetilde{h}^{(i)}(t) \triangleq h^{(i)}(T_{z_i}(t), T_{\mathcal{K}_i}(t)) - h^{(i)}(y^{(i)}(t), y_{\mathcal{K}_i}(t))$ and

$$\widetilde{g}^{(i)} = \sigma^{(i)}\left(n^{(i)} - n^s + f^{(i)} - f^s\right) \tag{24}$$

$$\begin{aligned}\widetilde{h}^{(i)} = {} & p^{(i)}\sum_{j\in\mathcal{K}_i} A_{d_{ij}}\left(\mu^{(i)}(T_{z_i}, T_{z_j}) - \mu^{(i)}(y^{(i)}, y^{(j)})\right) \\ & + \sum_{j\in\mathcal{K}_i}\frac{a_{z_{ij}}}{C_{z_i}}(n^{(j)} + f^{(j)})\end{aligned} \tag{25}$$

with

$$\mu^{(i)}(w_1, w_2) = \operatorname{sgn}(w_2 - w_1)\max(w_1, w_2)\sqrt{|w_2 - w_1|}. \tag{26}$$

Based on (11), the residual $\varepsilon^{(i)}$ defined in (21) can be re-written as

$$\varepsilon^{(i)}(t) = \varepsilon_T^{(i)}(t) + n^{(i)}(t) + f^{(i)}(t). \tag{27}$$

Remark 3: Given (10), (11), and (23)–(25), it yields that the residual $\varepsilon^{(i)}$ can be affected by a sensor fault in $\mathcal{S}^{(i)}$ and/or any sensor fault in $\mathcal{S}^{(j)}$, $j \in \mathcal{K}_i$ (sensors of neighboring subsystems) and/or sensor faults in $\mathcal{S}^s$.

2) Adaptive Threshold Design: The adaptive threshold $\bar{\varepsilon}^s(t) \in \mathbb{R}$ associated with the agent $\mathcal{M}^s$ is designed to bound the residual under healthy conditions (all sensors are healthy), which is denoted by $\varepsilon_H^s(t)$. Particularly, $\varepsilon_H^s(t)$ is defined as

$$\varepsilon_H^s(t) = \varepsilon_{T,H}^s(t) + n^s(t) \tag{28}$$

where $\varepsilon_{T,H}^s(t)$ is the state estimation error under healthy conditions that satisfies (17)–(19) with $f^s(t) = 0$ and $f^{(i)}(t) = 0$ for all $i \in \mathbb{N}$ and $y^s(t) = y_H^s(t)$, where

$$y_H^s(t) = T_{st}(t) + n^s(t). \tag{29}$$

Let us define $\bar{\varepsilon}_H^s(t) \in \mathbb{R}$ the adaptive threshold such that

$$|\varepsilon_H^s(t)| \le |\varepsilon_{T,H}^s(t)| + \bar{n}^s \le \bar{\varepsilon}_H^s(t),\ \forall\, t. \tag{30}$$

By bounding the solution of (17) and using (30), it yields

$$\begin{aligned}\bar{\varepsilon}_H^s(t) = {} & \bar{T}_{st}\rho^s e^{-\zeta^s t} + \bar{n}^s + \frac{\rho^s(1 - e^{-\zeta^s t})}{\zeta^s}(\bar{r}^s + \left|L^s\right|\bar{n}^s) \\ & + \int_0^t \rho^s e^{-\zeta^s(t-\tau)}\Big(\bar{g}^s(y_H^s(\tau))|u_{st}(\tau)| \\ & + \frac{a_{sz}}{C_{st}}\sum_{i\in\mathbb{N}} U_{i,\max}(\bar{n}^{(i)} + \bar{n}^s)|u_i(\tau)|\Big)d\tau\end{aligned} \tag{31}$$

where $\bar{T}_{st}$ is a bound on the initial state estimation error such that $\left|\varepsilon_T^s(0)\right| \le \bar{T}_{st}$, ρ^s, ζ^s are positive constants selected such

that $\left|e^{(A^s-L^s)t}\right| \le \rho^s e^{-\zeta^s t}$, for all t, and the function $\overline{g}^s(y_H^s)$ is defined in (70) and has been computed to bound $|\overline{g}^s|$ under healthy conditions as shown in the Appendix A.

Similarly, the adaptive threshold $\overline{\varepsilon}^{(i)}(t) \in \mathbb{R}$ associated with the agent $\mathcal{M}^{(i)}$ is designed to bound $|\varepsilon_H^{(i)}(t)|$ that denotes the residual under healthy conditions, defined as

$$\varepsilon_H^{(i)}(t) = \varepsilon_{T,H}^{(i)}(t) + n^{(i)}(t) \tag{32}$$

where $\varepsilon_{T,H}^{(i)}(t)$ is the state estimation error under healthy conditions that satisfies (23)–(25) with $f^s(t) = 0$, $f^{(i)}(t) = 0$, and $f^{(j)} = 0$, for all $j \in \mathcal{K}_i$ and $y^{(i)}(t) = y_H^{(i)}(t)$, where

$$y_H^{(i)}(t) = T_{z_i}(t) + n^{(i)}(t). \tag{33}$$

Let us define $\overline{\varepsilon}_H^{(i)}(t) \in \mathbb{R}$ the adaptive threshold under healthy conditions such that

$$|\varepsilon_H^{(i)}(t)| \le |\varepsilon_{T,H}^{(i)}(t)| + \overline{n}^{(i)} \le \overline{\varepsilon}_H^{(i)}(t),\ \forall\, t. \tag{34}$$

By bounding the solution of (23) and using (34), it yields

$$\begin{aligned}\overline{\varepsilon}_H^{(i)}(t) = {} & \overline{T}_{z_i}\rho^{(i)}e^{-\zeta^{(i)}t} + \overline{n}^{(i)} + \frac{\rho^{(i)}(1-e^{-\zeta^{(i)}t})}{\zeta^{(i)}}\Big(\left|L^{(i)}\right|\overline{n}^{(i)} + \overline{r}^{(i)} \\ & + \sum_{j\in\mathcal{K}_i}\frac{a_{z_i,j}}{C_{z_i}}\overline{n}^{(j)}\Big) + \int_0^t \rho^{(i)}e^{-\zeta^{(i)}(t-\tau)}\Big(\sigma^{(i)}\overline{n}^{(i)}|u^{(i)}(\tau)| \\ & + \sigma^{(i)}\overline{n}^{s}|u^{(i)}(\tau)| + p^{(i)}\sum_{j\in\mathcal{K}_i}A_{d_{i,j}}\overline{\mu}^{(i)}(y_H^{(i)}(\tau), y_H^{(j)}(\tau))\Big)d\tau\end{aligned} \tag{35}$$

where $\overline{T}_{z_i}$ is a bound such that $|\varepsilon_T^s(0)| \le \overline{T}_{z_i}$ and $\rho^{(i)}$, $\zeta^{(i)}$ are positive constants selected such that $|e^{\left(A^{(i)}-L^{(i)}\right)t}| \le \rho^{(i)}e^{-\zeta^{(i)}t}$, for all t. The function $\overline{\mu}^{(i)}$ is defined through (72)–(76) and (78)–(80) and is computed such that $|\mu^{(i)}(T_{z_i}, T_{z_j}) - \mu^{(i)}(y_H^{(i)}, y_H^{(j)})| \le \overline{\mu}^{(i)}(y_H^{(i)}, y_H^{(j)})$ as presented in Appendix B.

The adaptive thresholds $\overline{\varepsilon}^s(t)$ and $\overline{\varepsilon}^{(i)}(t)$ that are used for sensor fault detection are described by the following equations

$$\begin{aligned}\overline{\varepsilon}^s(t) = {} & \overline{T}_{st}\rho^s e^{-\zeta^s t} + \overline{n}^s + \frac{\rho^s(1-e^{-\zeta^s t})}{\zeta^s}(\overline{r}^s + \left|L^s\right|\overline{n}^s) \\ & + \int_0^t \rho^s e^{-\zeta^s(t-\tau)}\Big(\overline{g}^s(y^s(\tau))\,|u_{st}(\tau)| \\ & + \frac{a_{sz}}{C_{st}}\sum_{i\in\mathbb{N}}U_{i,\max}(\overline{n}^{(i)} + \overline{n}^s)\,|u_i(\tau)|\Big)d\tau\end{aligned} \tag{36}$$

$$\begin{aligned}\overline{\varepsilon}^{(i)}(t) = {} & \overline{T}_{z_i}\rho^{(i)}e^{-\zeta^{(i)}t} + \overline{n}^{(i)} + \frac{\rho^{(i)}(1-e^{-\zeta^{(i)}t})}{\zeta^{(i)}}\Big(\left|L^{(i)}\right|\overline{n}^{(i)} + \overline{r}^{(i)} \\ & + \sum_{j\in\mathcal{K}_i}\frac{a_{z_i,j}}{C_{z_i}}\overline{n}^{(j)}\Big) + \int_0^t \rho^{(i)}e^{-\zeta^{(i)}(t-\tau)}\Big(\sigma^{(i)}\overline{n}^{(i)}|u^{(i)}(\tau)| \\ & + \sigma^{(i)}\overline{n}^{s}|u^{(i)}(\tau)| + p^{(i)}\sum_{j\in\mathcal{K}_i}A_{d_{i,j}}\overline{\mu}^{(i)}(y^{(i)}(\tau), y^{(j)}(\tau))\Big)d\tau\end{aligned} \tag{37}$$

where y^s and $y^{(i)}$ are described by (10) and (11), respectively. The implementation of $\overline{\varepsilon}^s(t)$ and $\overline{\varepsilon}^{(i)}(t)$ can be realized using linear filtering techniques; i.e., $\int_0^t \rho e^{-\zeta(t-\tau)}z(\tau)d\tau$ can be implemented as $\frac{\rho}{s+\zeta}[z(t)]$ that corresponds to the output of the stable, linear filter $\rho/(s+\zeta)$ with input $z(t)$.

Remark 4: Note that under the occurrence of sensor faults, $\overline{\varepsilon}^s(t)$ may be affected by a fault in sensor $\mathcal{S}^s$ and $\overline{\varepsilon}^{(i)}(t)$ may be affected by faults in sensor $\mathcal{S}^{(i)}$ and $\mathcal{S}^{(j)}$ for all $j \in \mathcal{K}_i$.

3) Sensor Fault Detection Logic: The sensor fault detection process performed by the agents $\mathcal{M}^s$ and $\mathcal{M}^{(i)}$ is based on checking online whether the following analytical redundancy relations (ARR), denoted by $\mathcal{E}^s$ and $\mathcal{E}^{(i)}$, are satisfied

$$\mathcal{E}^s:\ \left|\varepsilon^s(t)\right| \le \overline{\varepsilon}^s(t) \tag{38}$$

$$\mathcal{E}^{(i)}:\ \left|\varepsilon^{(i)}(t)\right| \le \overline{\varepsilon}^{(i)}(t) \tag{39}$$

where ε^s and $\overline{\varepsilon}^s$ are defined in (15) and (36), while $\varepsilon^{(i)}$ and $\overline{\varepsilon}^{(i)}$ are given in (21) and (37). Hence, the boolean decision signal D^s (correspondingly $D^{(i)}$) indicates the violation of $\mathcal{E}^s$ (correspondingly of $\mathcal{E}^{(i)}$) such as D^s ($D^{(i)}$), i.e., when the threshold $\overline{\varepsilon}^s$ ($\overline{\varepsilon}^{(i)}$) is violated by the absolute value of the corresponding residual ε^s ($\varepsilon^{(i)}$, $i \in \mathbb{N}$).

B. Distributed Sensor Fault Isolation Module

When the detection decision signal D^s (correspondingly for $D^{(i)}$) becomes non-zero, the agent $\mathcal{M}^s$ (correspondingly for $\mathcal{M}^{(i)}$) initiates the fault isolation process, using local and neighboring detection decision signals.

The distributed isolation procedure applied by each agent involves the comparison of the observed pattern of sensor faults that may affect the neighborhood of the agent to a number of theoretical patterns, represented by the columns of a sensor fault signature matrix. In the case of the agent $\mathcal{M}^s$, the observed pattern of sensor faults, denoted by $\Phi^s(t) \in [0,1]^{N+1}$, where $[0,1]^{N+1}$ denotes a binary vector of $N+1$ length, defined as $\Phi^s(t) = [D^s, D^{(1)}, \ldots, D^{(N)}]$. Note that $D^{(i)}$ is transmitted to $\mathcal{M}^s$ by the agent $\mathcal{M}^{(i)}$ for all $i \in \mathbb{N}$. The sensor fault signature matrix F^s consists of $N+1$ rows, which correspond to the set of ARRs $\{\mathcal{E}^s, \mathcal{E}^{(1)}, \ldots, \mathcal{E}^{(N)}\}$, and $N_c = 2^{N+1} - 1$ columns that correspond to all possible sensor fault combinations that may affect the building zones and the storage tank, where the k-th combination is indicated by $\mathcal{F}_{c_k}^s$, $k \in \{1, \ldots, N_c\}$. The k-th column corresponds to the theoretical pattern, denoted by F_k^s and defined as $F_k^s = [F_{1k}^s, \ldots, F_{Nk}^s]^T$.

In the case of agent $\mathcal{M}^{(i)}$, the observed pattern of sensor faults, denoted by $\Phi^{(i)}(t) \in [0,1]^{|\mathcal{K}_i|+2}$, is a vector made up of the detection decisions $D^s(t)$, $D^{(i)}(t)$, and $D^{(j)}(t)$ for all $j \in \mathcal{K}_i$. The sensor fault signature matrix consists of $|\mathcal{K}_i| + 2$ rows, which correspond to the set of ARRs $\{\mathcal{E}^s, \mathcal{E}^{(i)}\} \bigcup_{j\in\mathcal{K}_i}\left\{\mathcal{E}^{(j)}\right\}$, and $N_c^{(i)} = 2^{|\mathcal{K}_i|+2} - 1$ columns that correspond to all possible sensor fault combinations that may affect the storage tank, the i-th building zone and its $|\mathcal{K}_i|$ neighboring zones. The k-th column corresponds to the theoretical pattern, denoted by $F_k^{(i)}$. For example, taking into account the 5-zone HVAC system shown in Fig. 2(c), based on which the observed pattern of agent $\mathcal{M}^s$ is defined as $\Phi^s(t) = \left[D^s(t), D^{(1)}(t), D^{(2)}(t), D^{(3)}(t), D^{(4)}(t), D^{(5)}(t)\right]$. Moreover, the sensor fault signature matrix F^s of the agent $\mathcal{M}^s$ presented in Fig. 2(c), is comprised of 6 rows and 63 columns as shown Table I, which illustrates a part of the sensor fault signature matrix F^s considering 6 single sensor faults, and one possible combination of two simultaneous sensor faults, $\{f^s, f^{(1)}\}$. The assignment $F_{22}^s = 1$

implies that $f^{(1)}$ necessarily discloses its occurrence by provoking the violation of $\mathcal{E}^{(1)}$, while $F^s_{12} = *$ implies that $f^{(1)}$ may justify the violation of $\mathcal{E}^s$, but $\mathcal{E}^s$ may be satisfied in spite of the occurrence of the sensor fault $f^{(1)}$. Otherwise, $F^s_{15} = 0$, since $f^{(5)}$ is not involved in $\mathcal{E}^{(1)}$ [43].

TABLE I
PART OF THE SENSOR FAULT SIGNATURE MATRIX OF THE AGENT $\mathcal{M}^s$ SHOWING IN FIG. 2(c)

	f^s	$f^{(1)}$	$f^{(2)}$	$f^{(3)}$	$f^{(4)}$	$f^{(5)}$	$\{f^s, f^{(1)}\}$
$\mathcal{E}^s$	1	*	*	*	*	*	1
$\mathcal{E}^{(1)}$	*	1	*	*	0	0	1
$\mathcal{E}^{(2)}$	*	*	1	*	0	*	*
$\mathcal{E}^{(3)}$	*	*	*	1	*	*	*
$\mathcal{E}^{(4)}$	*	0	0	*	1	*	*
$\mathcal{E}^{(5)}$	*	0	*	*	*	1	*

The sensor fault isolation process of the agent $\mathcal{M}^{(4)}$ presented in Fig. 2(c) consists of the observed pattern $\Phi^{(4)} = \left[D^s, D^{(4)}, D^{(3)}, D^{(5)}\right]$ and the sensor fault signature matrix $F^{(4)}$ comprised of 4 rows that corresponds to ARRs $\{\mathcal{E}^s, \mathcal{E}^{(4)}, \mathcal{E}^{(3)}, \mathcal{E}^{(5)}\}$ and 15 columns. Table II illustrates a part of $F^{(4)}$ considering 4 single sensor faults and 3 possible combinations of simultaneous sensor faults (i.e., $\{f^s, f^{(4)}\}$, $\{f^s, f^{(3)}\}$, and $\{f^s, f^{(5)}\}$).

TABLE II
PART OF THE SENSOR FAULT SIGNATURE MATRIX OF THE AGENT $\mathcal{M}^{(4)}$ SHOWING IN FIG. 2(c)

	f^s	$f^{(4)}$	$f^{(3)}$	$f^{(5)}$	$\{f^s, f^{(4)}\}$	$\{f^s, f^{(3)}\}$	$\{f^s, f^{(5)}\}$
$\mathcal{E}^s$	1	*	*	*	1	1	1
$\mathcal{E}^{(4)}$	*	1	*	*	1	*	*
$\mathcal{E}^{(3)}$	*	*	1	*	*	1	*
$\mathcal{E}^{(5)}$	*	*	*	1	*	*	1

Remark 5: The sensor fault isolation process of the agent $\mathcal{M}^{(4)}$ is realized in the neighborhood of $\mathcal{M}^{(4)}$ (see Fig. 2(c)) since the sensor faults $f^{(1)}$ and $f^{(2)}$ do not affect the residual generation of $\mathcal{M}^{(4)}$ (see (23)–(27) with $\mathcal{K}_4 = \{3, 5\}$).

The outcome of the online comparison of the observed fault pattern Φ^s to the N_c theoretical fault patterns F^s_k, $k \in \{1, \ldots, N_c\}$, and the observed pattern $\Phi^{(i)}$ to the $N^{(i)}_c$ theoretical patterns $F^{(i)}_q$, $q \in \{1, \ldots, N^{(i)}_c\}$ is the diagnosis sets $\Upsilon^s(t)$ and $\Upsilon^{(i)}(t)$, which are determined as

$$\Upsilon^s(t) = \left\{\mathcal{F}^s_{c_i} : i \in \mathcal{I}^s_\Upsilon(t)\right\}, \Upsilon^{(i)}(t) = \left\{\mathcal{F}^{(i)}_{c_i} : i \in \mathcal{I}^{(i)}_\Upsilon(t)\right\} \quad (40)$$

with $\mathcal{I}^s_\Upsilon(t) = \left\{k : F^s_k = \Phi^s(t),\ k \in \{1, \ldots, N_c\}\right\}$ and $\mathcal{I}^{(i)}_\Upsilon(t) = \left\{k : F^{(i)}_k = \Phi^{(i)}(t),\ k \in \{1, \ldots, N^{(i)}_c\}\right\}$. The diagnosis sets contains all the possible fault combinations.

IV. PERFORMANCE ANALYSIS

In this section we study the performance of the proposed sensor fault diagnosis architecture with respect to robustness (i.e., the ability to avoid false alarms in the presence of modeling uncertainty and measurement noise), detectability (i.e., the ability to detect faults in the presence of modeling uncertainty and measurement noise), and scalability (i.e., the ability to be easily modified in the case of increasing the number of zones).

A. Robustness Analysis

The property of robustness refers to the ability of the agents $\mathcal{M}^s$ and $\mathcal{M}^{(i)}$, $i \in \mathbb{N}$ to avoid false alarms in the presence of the modeling uncertainties r^s, $r^{(i)}$, and measurement noise n^s, $n^{(i)}$, in the absence of either local and propagated sensor fault. The robustness is accomplished by guaranteeing that the ARRs $\mathcal{E}^s$ and $\mathcal{E}^{(i)}$, respectively defined in (38) and (39), are satisfied, i.e., the magnitude of the residual remains below the adaptive threshold, under healthy conditions.

Lemma 1: If there are no faults affecting the sensor in the storage tank and all the sensors in the building zones, the ARR $\mathcal{E}^s$ is guaranteed to be satisfied and the agent $\mathcal{M}^s$ does not raise any false alarm in the presence of the modeling uncertainty r^s and measurement noise n^s and $n^{(i)}$ for all $i \in \mathbb{N}$.

Proof: If $f^s(t) = 0$ and $f^{(i)}(t) = 0$ for all $i \in \mathbb{N}$ then the residual $\varepsilon^s(t)$ is equal to $\varepsilon^s_H(t)$ defined in (28) and the adaptive threshold $\bar{\varepsilon}^s(t)$ is equal to $\bar{\varepsilon}^s_H(t)$ defined in (31). Therefore, (30) is valid and the ARR $\mathcal{E}^s$ defined in (38) is guaranteed to be satisfied. The robustness property is guaranteed based on the design of the fault diagnosis architecture. ■

Lemma 2: If there are no faults affecting the sensors in the storage tank and the building zone i, as well as the $|\mathcal{K}_i|$ sensors in the neighboring building zones, the ARR $\mathcal{E}^{(i)}$ is guaranteed to be satisfied and the agent $\mathcal{M}^{(i)}$ does not raise any false alarm in the presence of the modeling uncertainty $r^{(i)}$ and measurement noise n^s and $n^{(i)}$ for all $i \in \{\mathcal{K}_i \cup \{i\}\}$.

Proof: If $f^s(t) = 0$ and $f^{(i)}(t) = 0$ for all $i \in \{\mathcal{K}_i \cup \{i\}\}$ then the residual $\varepsilon^{(i)}(t)$ is equal to $\varepsilon^{(i)}_H(t)$ defined in (32) and the adaptive threshold $\bar{\varepsilon}^{(i)}(t)$ is equal to $\bar{\varepsilon}^{(i)}_H(t)$ defined in (35). Therefore, (34) is valid and the ARR $\mathcal{E}^{(i)}$ defined in (39) is guaranteed to be satisfied. ■

B. Detectability Analysis

This section contains the analysis on the detectability of the proposed distributed sensor fault diagnosis architecture where we analyze the ability of the agents to detect local and propagated sensor faults. Specifically, certain conditions are derived, under which we characterize the class of faults affecting the sensors in (10) and (11) that can be detected. It is important to note that the class of detectable sensor faults satisfying these conditions is obtained under worst-case assumptions, in the sense that they are valid for any modeling uncertainty and measurement noise satisfying Assumption 1. The analysis is divided into two parts; the sensor fault detectability analysis of agent $\mathcal{M}^s$ and the sensor fault detectability analysis of agent $\mathcal{M}^{(i)}$, $i \in \mathbb{N}$.

1) Sensor Fault Detectability of Agent $\mathcal{M}^s$: The residual ε^s described by (15) (or (20)) and the corresponding adaptive threshold $\bar{\varepsilon}^s$ of (36) are sensitive to any fault that may occur in the sensor of the storage tank (local sensor fault) at the time instant t^s_f, or in the sensors of the building zones (propagated sensor faults) that may occur at the time instances $t^{(i)}_f$, $i \in \mathbb{N}$. Under faulty conditions, ε^s and $\bar{\varepsilon}^s$ can be expressed as

$$\varepsilon^s(t) = \varepsilon^s_H(t) + \varepsilon^s_F(t) \tag{41}$$

$$\overline{\varepsilon}^s(t) = \overline{\varepsilon}^s_H(t) + \overline{\varepsilon}^s_F(t) \tag{42}$$

where ε^s_H (defined in (20)) and $\overline{\varepsilon}^s_H$ (defined in (31)) are the healthy parts of ε^s and $\overline{\varepsilon}^s$, respectively, and ε^s_F and $\overline{\varepsilon}^s_F$ are the faulty parts of ε^s and $\overline{\varepsilon}^s$, respectively, which include the effects of faults [43]. Given (30), (41), and (42), sensor faults are guaranteed to be detected if there exists a $t^\star$ such that

$$\left|\varepsilon^s_F(t^\star)\right| - \overline{\varepsilon}^s_F(t^\star) > 2\overline{\varepsilon}^s_H(t^\star). \tag{43}$$

Condition (43) guarantees the violation of ARR $\mathcal{E}^s$ given in (38). The sensor fault effects ε^s_F and $\overline{\varepsilon}^s_F$ can be characterized taking into account the occurrence of

1) a local sensor fault $f^s(t)$ for $t \in [t^s_f, \min_{i\in\mathbb{N}}\{\mathrm{t}^{(i)}_f\})$;

2) propagated sensor faults $f^{(i)}(t)$ for $t \in [\min_{i\in\mathbb{N}}\{\mathrm{t}^{(i)}_f\}, t^s_f)$ with $\max_{i\in\mathbb{N}}\{\mathrm{t}^{(i)}_f\} < t^s_f$;

3) both local $f^s(t)$ and propagated sensor faults $f^{(i)}(t)$ for $t \geq \max(t^s_f, \max_{i\in\mathbb{N}}\{\mathrm{t}^{(i)}_f\})$.

Lemma 3: A sensor fault f^s affecting the temperature sensor $\mathcal{S}^s$ at the time instant t^s_f is guaranteed to be detected by $\mathcal{M}^s$, if there exists a time instant $t^\star \in (t^s_f, \min_{i\in\mathbb{N}}\{\mathrm{t}^{(i)}_f\})$ such that

$$\begin{aligned} 2\overline{\varepsilon}^s_H(t^\star) \; < \; & \Bigg| f^s(t^\star) + \int_{t^s_f}^{t^\star} e^{(A^s-L^s)(t^\star-\tau)}\Big(-L^s f^s(\tau) \\ & + \frac{U_{st,\max}}{C_{st}}\Big(P_s(y^s_H(\tau)) - P_s(y^s_H(\tau)+f^s(\tau))\Big)u_{st}(\tau) \\ & - \frac{a_{sz}}{C_{st}}\sum_{j\in\mathbb{N}} U_{i,\max}u_i(\tau)f^s(\tau)\Big)d\tau \Bigg| - \int_{t^s_f}^{t^\star} \rho^s e^{-\zeta^s(t-\tau)} \\ & \times\Big(\Big(\overline{g}^s(y^s_H(\tau)+f^s(\tau)) - \overline{g}^s(y^s_H(\tau))\Big)|u_{st}(\tau)|\Big)d\tau \end{aligned} \tag{44}$$

where y^s_H is defined in (29).

Proof: Under healthy conditions the residual ε^s equals to ε^s_H defined in (28), where the state estimation error under healthy conditions $\varepsilon^s_{T,H}$ corresponds to the solution of (17), taking into account that $f^s(t) = 0$ and $f^{(i)}(t) = 0$ for all $i \in \mathbb{N}$ and $y^s(t) = y^s_H(t)$ where y^s_H is defined in (29); i.e.,

$$\begin{aligned} \varepsilon^s_{T,H}(t) = & \varepsilon^s_{T,H}(0)e^{(A^s-L^s)t} + \int_0^t e^{(A^s-L^s)(t-\tau)}\Big(r^s(\tau) \\ & + \frac{U_{st,\max}}{C_{st}}\Big(P_s(T_{st}(\tau)) - P_s(y^s_H(\tau))\Big)u_{st}(\tau) - L^s n^s(\tau) \\ & + \frac{a_{sz}}{C_{st}}\sum_{i\in\mathbb{N}} U_{i,\max}(n^{(i)}(\tau) - n^s(\tau))u^{(i)}(\tau)\Big)d\tau \end{aligned} \tag{45}$$

where $\varepsilon^s_{T,H}(0) = \varepsilon^s_T(0)$. Assuming local sensor fault ($f^s(t) \neq 0$ and $f^{(i)}(t) = 0$ for all $i \in \mathbb{N}$), the state estimation error is given by the solution of (17) for $t \geq t^s_f$; i.e.,

$$\begin{aligned} \varepsilon^s_T(t) = & \varepsilon^s_T(t^s_f)e^{(A^s-L^s)\left(t-t^s_f\right)} + \int_{t^s_f}^t e^{(A^s-L^s)(t-\tau)}\Big(r^s(\tau) \\ & + \frac{U_{st,\max}}{C_{st}}\Big(P_s(T_{st}(\tau)) - P_s(y^s_H(\tau)+f^s(\tau))\Big)u_{st}(\tau) \\ & + \frac{a_{sz}}{C_{st}}\sum_{i\in\mathbb{N}} U_{i,\max}(n^{(i)}(\tau) - n^s(\tau) - f^s(\tau))u_i(\tau) \\ & - L^s n^s(\tau) - L^s f^s(\tau)\Big)d\tau. \end{aligned} \tag{46}$$

Based on (12) and (14), $\varepsilon^s_T(t^s_f) = \varepsilon^s_{T,H}(t^s_f)$. By using (45) for $t = t^s_f$ in (46) it yields

$$\begin{aligned} \varepsilon^s_T(t) = & \varepsilon^s_{T,H}(t) + \int_{t^s_f}^t e^{(A^s-L^s)(t-\tau)}\Big(-L^s f^s(\tau) \\ & + \frac{U_{st,\max}}{C_{st}}\Big(P_s(y^s_H(\tau)) - P_s(y^s_H(\tau)+f^s(\tau))\Big)u_{st}(\tau) \\ & - \frac{a_{sz}}{C_{st}}\sum_{i\in\mathbb{N}} U_{i,\max}f^s(\tau)u_i(\tau)\Big)d\tau. \end{aligned} \tag{47}$$

Combining (20), (28), and (41) result in

$$\varepsilon^s_F(t) = \varepsilon^s(t) - \varepsilon^s_H(t) = \varepsilon^s_T(t) - \varepsilon^s_{T,H}(t) + f^s(t). \tag{48}$$

By introducing (47) in (48), we obtain

$$\begin{aligned} \varepsilon^s_F(t) = & f^s(t) + \int_{t^s_f}^t e^{(A^s-L^s)(t-\tau)}\Big(-L^s f^s(\tau) \\ & + \frac{U_{st,\max}}{C_{st}}\Big(P_s(y^s_H(\tau)) - P_s(y^s_H(\tau)+f^s(\tau))\Big)u_{st}(\tau) \\ & - \frac{a_{sz}}{C_{st}}\sum_{i\in\mathbb{N}} U_{i,\max}f^s(\tau)u_i(\tau)\Big)d\tau. \end{aligned} \tag{49}$$

The effects of sensor faults on the adaptive threshold can be determined using (31), (36), and (42) as

$$\begin{aligned} \overline{\varepsilon}^s_F(t) = \overline{\varepsilon}^s(t) - \overline{\varepsilon}^s_H(t) = & \int_0^t \rho^s e^{-\zeta^s(t-\tau)}\Big(\overline{g}^s(y^s_H(\tau)+f^s(\tau)) \\ & - \overline{g}^s(y^s_H(\tau))\Big)|u_{st}(\tau)|\,d\tau. \end{aligned} \tag{50}$$

Based on (12) and (14), $f^s(t) = 0$ for $t < t^s_f$ implying that $\int_0^{t^s_f} \rho^s e^{-\zeta^s(t-\tau)}\Big(\overline{g}^s(y^s_H(\tau)+f^s(\tau)) - \overline{g}^s(y^s_H(\tau))\Big)|u_{st}(\tau)|\,d\tau = 0$. Thus (50) becomes

$$\begin{aligned} \overline{\varepsilon}^s_F(t) = & \int_{t^s_f}^t \rho^s e^{-\zeta^s(t-\tau)} \\ & \times(\overline{g}^s(y^s_H(\tau)+f^s(\tau)) - \overline{g}^s(y^s_H(\tau)))\,|u_{st}(\tau)|\,d\tau. \end{aligned} \tag{51}$$

Introducing (49) and (51) in (43) leads to (44). ■

The conditions for guaranteeing the detection of (possibly multiple) propagated faults that affect the sensors located in the building zones by the agent $\mathcal{M}^s$ is analyzed in Lemma 4. It is worth noting that the propagated sensor faults $f^{(i)}$ can affect the residual ε^s defined through (17)–(20) and not the adaptive thresholds $\overline{\varepsilon}^s(t)$ defined in (36).

Lemma 4: Sensor faults $f^{(i)}$ affecting the temperature sensors $\mathcal{S}^{(i)}$ in the building zones at the time instances $t^{(i)}_f$ are guaranteed to be detected by $\mathcal{M}^s$, if there exists a time instant $t^\star \in (\min_{i\in\mathbb{N}}\{\mathrm{t}^{(i)}_f\}, t^s_f)$ with $\max_{i\in\mathbb{N}}\{\mathrm{t}^{(i)}_f\} < t^s_f$ such that

$$2\overline{\varepsilon}^s_H(t^\star) < \left|\frac{a_{sz}}{C_{st}}\sum_{i\in\mathbb{N}}\int_{t^{(i)}_f}^{t^\star} e^{(A^s-L^s)(t^\star-\tau)}U_{i,\max}f^{(i)}(\tau)u_i(\tau)d\tau\right|. \tag{52}$$

Proof: Under healthy conditions the residual ε^s equals to ε^s_H defined in (28), where the state estimation error under healthy conditions $\varepsilon^s_{T,H}$ is defined in (45). Let us consider two propagated sensor faults in, e.g., zones 1 and 2, where sensor fault $f^{(1)}$ occurs at $t^{(1)}_f$ and sensor fault $f^{(2)}$ occurs at $t^{(2)}_f$ with $t^{(1)}_f < t^{(2)}_f$. Based on the state estimation error dynamics given in (17), $\varepsilon^s_T(t)$ for $t \in [t^{(1)}_f, t^{(2)}_f)$ is given by

$$\varepsilon_T^s(t) = \varepsilon_T^s(t_f^{(1)}) e^{(A^s - L^s)\left(t - t_f^{(1)}\right)} + \int_{t_f^{(1)}}^{t} e^{(A^s - L^s)(t-\tau)} \Big(r^s(\tau) + \frac{U_{st,\max}}{C_{st}} \left(P_s(T_{st}(\tau)) - P_s(y_H^s(\tau)) \right) u_{st}(\tau) - L^s n^s(\tau) + \frac{a_{sz}}{C_{st}} \sum_{i \in \mathbb{N}} U_{i,\max} (n^{(i)}(\tau) - n^s(\tau)) u_i(\tau) + \frac{a_{sz}}{C_{st}} U_{1,\max} f^{(1)}(\tau) u_1(\tau) \Big) d\tau \tag{53}$$

while for $t \geq t_f^{(2)}$ $\varepsilon_T^s(t)$ is expressed as

$$\varepsilon_T^s(t) = \varepsilon_T^s(t_f^{(2)}) e^{(A^s - L^s)\left(t - t_f^{(2)}\right)} + \int_{t_f^{(2)}}^{t} e^{(A^s - L^s)(t-\tau)} \Big(r^s(\tau) + \frac{U_{st,\max}}{C_{st}} \left(P_s(T_{st}(\tau)) - P_s(y_H^s(\tau)) \right) u_{st}(\tau) + \frac{a_{sz}}{C_{st}} \sum_{i \in \mathbb{N}} U_{i,\max} (n^{(i)}(\tau) - n^s(\tau)) u_i(\tau) - L^s n^s(\tau) + \frac{a_{sz}}{C_{st}} (U_{1,\max} f^{(1)}(\tau) u_1(\tau) + U_{2,\max} f^{(2)}(\tau) u_2(\tau)) \Big) d\tau. \tag{54}$$

By using (45) for $t = t_f^{(2)}$ in (53), and then using (53) for $t = t_f^{(2)}$ in (54) it yields

$$\varepsilon_T^s(t) = \varepsilon_{T,H}^s(t) + \frac{a_{sz}}{C_{st}} \Bigg(\int_{t_f^{(1)}}^{t} e^{(A^s - L^s)(t-\tau)} U_{1,\max} f^{(1)}(\tau) u_1(\tau) d\tau + \int_{t_f^{(2)}}^{t} e^{(A^s - L^s)(t-\tau)} U_{2,\max} f^{(2)}(\tau) u_2(\tau) d\tau \Bigg). \tag{55}$$

Equation (55) is also valid in the case that $t_f^{(2)} < t_f^{(1)}$. If we perform the same mathematical manipulations, we can obtain that the state estimation error $\varepsilon_T^s(t)$ for $t \in (\min_{i \in \mathbb{N}} \{\mathrm{t}_f^{(i)}\}, t_f^s)$ with $\max_{i \in \mathbb{N}} \{\mathrm{t}_f^{(i)}\} < t_f^s$ is described by

$$\varepsilon_T^s(t) = \varepsilon_{T,H}^s(t) + \frac{a_{sz}}{C_{st}} \sum_{i \in \mathbb{N}} \int_{t_f^{(i)}}^{t} e^{(A^s - L^s)(t-\tau)} U_{i,\max} f^{(i)}(\tau) u_i(\tau) d\tau. \tag{56}$$

By combining (48) with $f^s = 0$ and (56), the effects of propagated sensor faults on the residual are described by

$$\varepsilon_F^s(t) = \frac{a_{sz}}{C_{st}} \sum_{i \in \mathbb{N}} \int_{t_f^{(i)}}^{t} e^{(A^s - L^s)(t-\tau)} U_{i,\max} f^{(i)}(\tau) u_i(\tau) d\tau. \tag{57}$$

Using (57) in (43) and given that $\bar{\varepsilon}_F^s(t) = 0$ leads to (52). ■

Lemma 5: The sensor faults f^s and $f^{(i)}$ that occur at the time instants t_f^s and $t_f^{(i)}$, respectively are guaranteed to be detected by $\mathcal{M}^s$, if there exist a time instant $t^\star \geq \max(t_f^s, \max_{i \in \mathbb{N}} \{\mathrm{t}_f^{(i)}\})$ such that

$$2\bar{\varepsilon}_H^s(t^\star) < \Bigg| f^s(t^\star) + \int_{\max(t_f^s, \max_{i \in \mathbb{N}} \{\mathrm{t}_f^{(i)}\})}^{t^\star} e^{(A^s - L^s)(t^\star - \tau)} \Bigg(- L^s f^s(\tau) + \frac{U_{st,\max}(P_{\max} - 1)}{C_{st} \Delta T_{\max}} u_{st}(\tau) f^s(\tau) + \frac{a_{sz}}{C_{st}} \sum_{j \in \mathbb{N}} U_{i,\max} u_i(\tau) \left(f^{(i)}(\tau) - f^s(\tau) \right) \Bigg) d\tau \Bigg| - \int_{\max(t_f^s, \max_{i \in \mathbb{N}} \{\mathrm{t}_f^{(i)}\})}^{t^\star} \rho^s e^{-\zeta^s (t-\tau)} \Big((\bar{g}^s(y_H^s(\tau) + f^s(\tau)) - \bar{g}^s(y_H^s(\tau))) |u_{st}(\tau)| \Big) d\tau. \tag{58}$$

The proof of Lemma 5 is not provided, but it can be obtained similarly as in Lemmas 3 and 4. Lemmas 3–5 provide certain conditions that characterize analytically the class of local and propagated sensor faults that are guaranteed to be detectable by the agent $\mathcal{M}^s$.

2) Sensor Fault Detectability Analysis of Agent $\mathcal{M}^{(i)}$: The residual $\varepsilon^{(i)}$ given in (21) (or (27)) and the corresponding adaptive threshold $\bar{\varepsilon}^{(i)}$ of (37) are sensitive to any faults that may occur in the building zone i (local sensor fault) at the time instant $t_f^{(i)}$, or in the sensor of the storage tank at the time instant t_f^s, or in the $|\mathcal{K}_i|$ neighboring zones (propagated sensor faults) that may occur at the time instances $t_f^{(j)}$, $j \in \mathcal{K}_i$. Under faulty conditions, $\varepsilon^{(i)}$ and $\bar{\varepsilon}^{(i)}$ can be expressed as

$$\varepsilon^{(i)}(t) = \varepsilon_H^{(i)}(t) + \varepsilon_F^{(i)}(t) \tag{59}$$

$$\bar{\varepsilon}^{(i)}(t) = \bar{\varepsilon}_H^{(i)}(t) + \bar{\varepsilon}_F^{(i)}(t) \tag{60}$$

where $\varepsilon_H^{(i)}$ (defined in (27)) and $\bar{\varepsilon}_H^{(i)}$ (defined in (32)) are the healthy parts of $\varepsilon^{(i)}$ and $\bar{\varepsilon}^{(i)}$, respectively, and $\varepsilon_F^{(i)}$ and $\bar{\varepsilon}_F^{(i)}$ are the faulty parts of $\varepsilon^{(i)}$ and $\bar{\varepsilon}^{(i)}$, which include the effects of faults. Given (34), (59), and (60), sensor faults are guaranteed to be detected if there exists a time instant $t^\star$ such that

$$\left| \varepsilon_F^{(i)}(t^\star) \right| - \bar{\varepsilon}_F^{(i)}(t^\star) > 2\bar{\varepsilon}_H^{(i)}(t^\star). \tag{61}$$

Condition (61) guarantees the violation of ARR $\mathcal{E}^{(i)}$ given in (39). The sensor fault effects $\varepsilon_F^{(i)}$ and $\bar{\varepsilon}_F^{(i)}$ can be characterized taking into account the occurrence of

1) a local sensor fault $f^{(i)}(t)$ for $t \in [t_f^{(i)}, \min(\min_{j \in \mathcal{K}_i} \{\mathrm{t}_f^{(i)}\}, t_f^s))$;

2) propagated sensor faults $f^{(j)}(t)$ for $t \in [\min_{j \in \mathcal{K}_i} \{\mathrm{t}_f^{(i)}\}, t_f^s)$ with $\max(\max_{j \in \mathcal{K}_i} \{\mathrm{t}_f^{(j)}\}, t_f^s) < t_f^{(i)}$;

3) both local $f^{(i)}(t)$ and propagated sensor faults $f^s(t)$, $f^{(j)}(t)$ for $t \geq \max(t_f^{(j)}, t_f^s, \max_{i \in \mathbb{N}} \{\mathrm{t}_f^{(i)}\})$.

The proofs of the following Lemmas 6–8 are not given, but they can be obtained similarly as the proofs of Lemmas 3 and 4.

Lemma 6: The sensor fault $f^{(i)}$ affecting the temperature sensor $\mathcal{S}^{(i)}$ at the time instant $t_f^{(i)}$ is guaranteed to be detected by $\mathcal{M}^{(i)}$ under worst-case conditions, if there exist a time instant $t^\star \in [t_f^{(i)}, \min(\min_{j \in \mathcal{K}_i} \{\mathrm{t}_f^{(i)}\}, t_f^s))$ such that

$$2\bar{\varepsilon}_H^{(i)}(t^\star) < \Bigg| f^{(i)}(t^\star) + \int_{t_f^{(i)}}^{t^\star} e^{\left(A^{(i)}-L^{(i)}\right)(t-\tau)} \Bigg(\sigma^{(i)} f^{(i)}(\tau) u_i(\tau) + p^{(i)} \sum_{j\in\mathcal{K}_i} A_{d_{ij}} \left(\mu^{(i)}(y_H^{(i)}(\tau), y_H^{(j)}(\tau)) - \mu^{(i)}(y_H^{(i)}(\tau) + f^{(i)}(\tau), y_H^{(j)}(\tau)) \right) - L^{(i)} f^{(i)}(\tau) \Bigg) d\tau \Bigg| - \int_{t_f^{(i)}}^{t^\star} \rho^{(i)} e^{-\zeta^{(i)}(t-\tau)} p^{(i)} \sum_{j\in\mathcal{K}_i} A_{d_{i,j}} \times \left(\bar{\mu}^{(i)}(y_H^{(i)}(\tau) + f^{(i)}(\tau), y_H^{(j)}(\tau)) - \bar{\mu}^{(i)}(y_H^{(i)}(\tau), y_H^{(j)}(\tau)) \right) d\tau. \quad (62)$$

Lemma 7: The sensor faults f^s and $f^{(j)}$ occur at the time instants t_f^s and $t_f^{(j)}$, respectively are guaranteed to be detected by $\mathcal{M}^{(i)}$ under worst-case conditions, if there exist a time instant $t^\star \in [\min_{j\in\mathcal{K}_i}\{\mathrm{t}_f^{(i)}\}, t_f^s)$ with $\max(\max_{j\in\mathcal{K}_i}\{\mathrm{t}_f^{(j)}\}, t_f^s) < t_f^{(i)}$ such that

$$2\bar{\varepsilon}_H^{(i)}(t^\star) < \Bigg| \int_{\min_{j\in\mathcal{K}_i}\{\mathrm{t}_f^{(i)}\}}^{t^\star} e^{\left(A^{(i)}-L^{(i)}\right)(t-\tau)} \Bigg(-\sigma^{(i)} f^s(\tau) u_i(\tau) + p^{(i)} \sum_{j\in\mathcal{K}_i} A_{d_{ij}} \left(\mu^{(i)}(y_H^{(i)}(\tau), y_H^{(j)}(\tau)) - \mu^{(i)}(y_H^{(i)}(\tau), y_H^{(j)}(\tau) + f^{(j)}(\tau)) \right) + \sum_{j\in\mathcal{K}_i} \frac{a_{z_{ij}}}{C_{z_i}} f^{(j)}(\tau) \Bigg) d\tau \Bigg| - \int_{\min_{j\in\mathcal{K}_i}\{\mathrm{t}_f^{(i)}\}}^{t^\star} \rho^{(i)} e^{-\zeta^{(i)}(t-\tau)} p^{(i)} \times \sum_{j\in\mathcal{K}_i} A_{d_{i,j}} \left(\bar{\mu}^{(i)}(y_H^{(i)}(\tau), y_H^{(j)}(\tau) + f^{(j)}(\tau)) - \bar{\mu}^{(i)}(y_H^{(i)}(\tau), y_H^{(j)}(\tau)) \right) d\tau. \quad (63)$$

Lemma 8: The sensor faults $f^{(i)}$, f^s, and $f^{(j)}$ occur at the time instants $t_f^{(i)}$, t_f^s, and $t_f^{(j)}$, respectively are guaranteed to be detected by $\mathcal{M}^{(i)}$ under worst-case conditions, if there exist a time instant $t^\star \geq \max(t_f^{(j)}, t_f^s, \max_{i\in\mathbb{N}}\{\mathrm{t}_f^{(i)}\})$ such that

$$2\bar{\varepsilon}_H^{(i)}(t^\star) < \Bigg| f^{(i)}(t^\star) + \int_{\max(t_f^{(j)}, t_f^s, \max_{i\in\mathbb{N}}\{\mathrm{t}_f^{(i)}\})}^{t^\star} e^{\left(A^{(i)}-L^{(i)}\right)(t-\tau)} \times \Bigg(\sigma^{(i)}(f^{(i)}(\tau) - f^s(\tau)) u_i(\tau) + p^{(i)} \sum_{j\in\mathcal{K}_i} A_{d_{ij}} \times \left(\mu^{(i)}(y_H^{(i)}(\tau), y_H^{(j)}(\tau)) - \mu^{(i)}(y_H^{(i)}(\tau) + f^{(i)}(\tau), y_H^{(j)}(\tau) + f^{(j)}(\tau)) \right) + \sum_{j\in\mathcal{K}_i} \frac{a_{z_{ij}}}{C_{z_i}} f^{(j)}(\tau) - L^{(i)} f^{(i)}(\tau) \Bigg) d\tau \Bigg| - \int_{\max(t_f^{(j)}, t_f^s, \max_{i\in\mathbb{N}}\{\mathrm{t}_f^{(i)}\})}^{t^\star} \rho^{(i)} e^{-\zeta^{(i)}(t-\tau)} p^{(i)} \sum_{j\in\mathcal{K}_i} A_{d_{i,j}} \times \left(\bar{\mu}^{(i)}(y_H^{(i)}(\tau) + f^{(i)}(\tau), y_H^{(j)}(\tau) + f^{(j)}(\tau)) - \bar{\mu}^{(i)}(y_H^{(i)}(\tau), y_H^{(j)}(\tau)) \right) d\tau. \quad (64)$$

Lemmas 6–8 provide certain conditions that characterize analytically the class of local and propagated sensor faults that are guaranteed to be detectable by the agent $\mathcal{M}^{(i)}$.

Remark 6: The detectability conditions obtained in Lemmas 3–5 and 6–8, give an indication of the class of local and propagated sensor faults that are guaranteed to be detectable by the agents $\mathcal{M}^s$ and $\mathcal{M}^{(i)}$, respectively based on the system's available parameters and proposed algorithm's design parameters. However, due to the non-linear and switching terms in systems' dynamics, the aforementioned conditions depend on real-time signals; thus, obtaining off-line predefined, fixed conditions is not possible.

The above issue can be addressed by creating a Monte-Carlo analysis, examining the detectability performance by varying the sensor noise, modeling uncertainty, and observer design parameters (see simulation-based analysis presented in the Section V.)

C. Scalability Analysis

This subsection provides a discussion on the scalability of the proposed distributed sensor fault diagnosis technique in the case that the multi-zone HVAC system is enlarged with respect to the number of building zones. For example, a new building zone may be constructed, whose temperature is monitored by a sensor and controlled by a fan-coil unit. In the following analysis we consider the aforementioned example. A similar discussion can be considered for the case that some buildings zones are removed.

Consider that a 6-th zone is constructed next to the 5-th zone of the HVAC system shown in Fig. 2(a), while there is a door (and walls) connecting the two zones as shown in Fig. 3(a). The 6-th zone is comprised of a temperature sensor and a fan-coil unit connected to the central electromechanical part. Given the architectural/thermal parameters and the manufacturing properties of the fan-coil unit installed in the new zone, the subsystem $\Sigma^{(6)}$ (green box in Fig. 3(b)) is defined according to (7) with $i = 6$ and $\mathcal{K}_6 = \{5\}$. The equations in Table III describe the modification of the existing HVAC model according to the physical variation of the HVAC system for $t < t_{en}$, where t_{en} is the time at which the HVAC system is enlarged. Note that χ^s and $\chi^{(5)}$ collect the dynamic terms of the electromechanical part and the 5-th zone, respectively. The agent $\mathcal{M}^{(6)}$ is designed based on (21), (22), (37), and (39). Only the agents $\mathcal{M}^s$ and $\mathcal{M}^{(5)}$, presented with purple boxes in Fig. 3(c), should be modified based on a plugin process shown in Table IV. The existing estimator and adaptive threshold of these agents ($\mathcal{M}^s$ and $\mathcal{M}^{(5)}$) are not modified but some new plug-in blocks are added. For agent $\mathcal{M}^{(5)}$, the plug-in blocks are illustrated with gray color in Fig. 4. This allows to scale the sensor fault diagnosis scheme even without re-designing any agents of the 5-th zone HVAC system. The scalability property, which is possible due to the distributed monitoring architecture of the proposed scheme, is important in large-scale systems since it allows the evolution of the HVAC with additional zones without having to redesign the overall system.

V. Simulation Results

The objective of this section is the evaluation of the

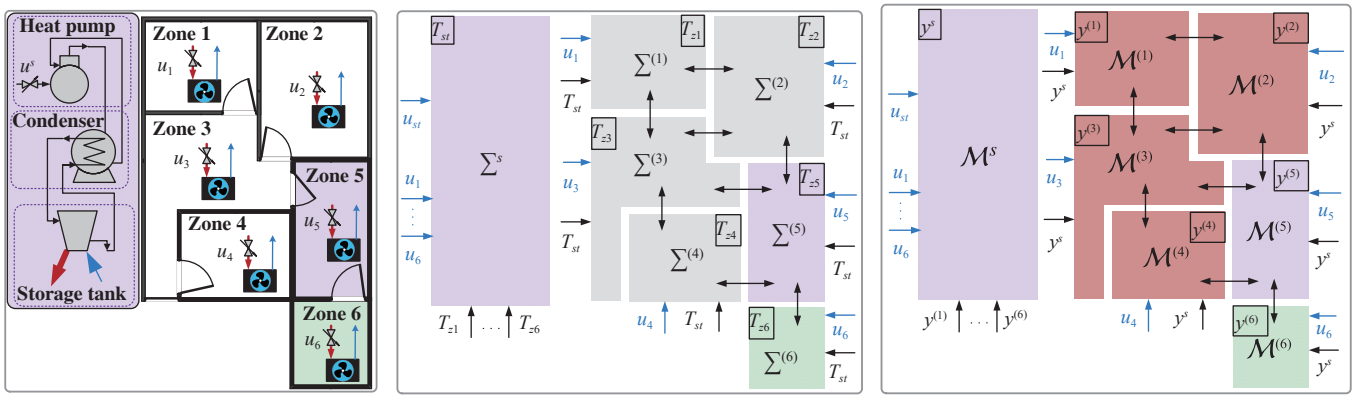


Fig. 3. Reconfiguration of the distributed sensor fault diagnosis architecture for the enlarged HVAC system. The 6-th zone (green floor) is added to (a) which is connected to the 1-st zone; In (b) and (c) the reconfiguration of the network of interconnected subsystems and the reconfiguration of the sensor fault diagnosis agents are presented, respectively. Green color denotes the added components/subsystems/agents while the purple color denotes the modified components/subsystems/agents.

TABLE III
MODEL VARIATIONS AFTER THE ENLARGEMENT OF THE HVAC SYSTEM

Σ^s	$\dot{T}_{st} = \chi^s(T_{st}, T_z, u) + \frac{a_{sz}}{C_{st}} U_{6,\max}(T_{st} - T_{z_6})u_6$
$\Sigma^{(5)}$, $\mathcal{K}_5 = \{2,3,4\}$	$\dot{T}_{z_5} = \chi^{(5)}(T_{z_5}, T_{st}, T_{\mathcal{K}_5}, u_5) + a_{z_{5,6}} T_{z_6}/C_{z_1}$ $+ p^{(5)} A_{d_{5,6}} \mu^{(5)}(T_{z_5}, T_{z_6})$

TABLE IV
DESIGN PLUG-IN BLOCKS TO THE SENSOR FAULT DIAGNOSIS SCHEME

$\mathcal{M}^s$	$\mathcal{E}^s: \ \lvert\varepsilon^s\rvert \le \bar{\varepsilon}^s + \bar{\varepsilon}^{s(6)}$
	$\varepsilon^s = y^s - \hat{T}_{st} - \hat{T}_{st(6)}$
	$\dot{\hat{T}}_{st(6)} = (A^s - L^s)\hat{T}_{st(6)} + \frac{a_{sz}}{C_{st}} U_{6,\max}(y^s - y^{(6)})u_6,$
	$\hat{T}_{st(6)}(t_{en}) = 0$
	$\bar{\varepsilon}^{s(6)} = \int_{t_e}^{t} \rho^s e^{-\zeta^s(t-\tau)} \frac{a_{sz}}{C_{st}} U_{6,\max}(\bar{n}^s + \bar{n}^{(6)})\lvert u_6(\tau)\rvert$
$\mathcal{M}^{(5)}$	$\mathcal{E}^{(5)}: \ \lvert\varepsilon^{(5)}\rvert \le \bar{\varepsilon}^{(5)} + \bar{\varepsilon}^{(5,6)}$
	$\varepsilon^{(5)} = y^{(5)} - \hat{T}_{z_5} - \hat{T}_{5,6}$
	$\dot{\hat{T}}_{5,6} = \left(A^{(5)} - L^{(5)}\right)\hat{T}_{5,6} + \frac{1}{C_{z_5}} a_{z_{5,6}} y^{(6)}$ $+ p^{(5)} A_{d_{5,6}} \mu^{(5)}(y^{(5)}, y^{(6)}),$
	$\hat{T}_{5,6}(t_{en}) = 0$
	$\bar{\varepsilon}^{(5,6)} = \int_{t_e}^{t} \rho^{(5)} e^{-\zeta^{(5)}(t-\tau)} p^{(5)} A_{d_{5,6}} \bar{\mu}^{(5)}(y^{(5)}, y^{(6)})$ $+ \frac{\rho^{(5)}(1 - e^{-\zeta^{(5)}t})}{\zeta^{(5)}} \frac{a_{z_{5,6}}}{C_{z_5}} \bar{n}^{(6)}$

proposed distributed fault diagnosis method applied to a large-scale building. Let us consider a 83-zone HVAC system whose down-view is presented in Fig. 5. Table V provides a list of parameters for the 83-zone HVAC system. As shown in Fig. 5 the building consists of 16 apartments (5-zones each), 2 stair halls and 1 corridor. The structural properties of each apartment are the same, hence the Table V contains the parameters of one apartment (i.e., zones 1–5), one stair hall (i.e., zone 81), and the corridor (i.e., zone 83). The remainder parameters of the 83-zone HVAC system are: $a_{st} = 12$ kJ/kg °C, $a_{sz} = 0.6$ kJ/kg °C, $U_{st,\max} = 27.36 \times 10^5$ kJ/°C,

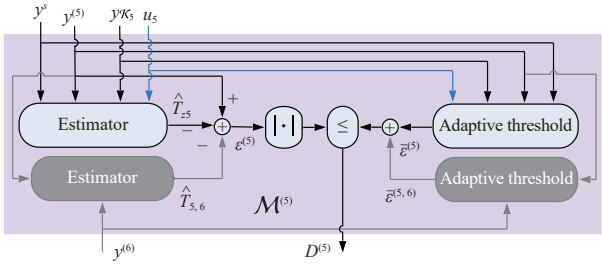


Fig. 4. The gray boxes and arrows denote the plugin blocks and signals added to the existing agent $\mathcal{M}^{(5)}$ at $t \ge t_{en}$, with $y_{\mathcal{K}_5} = \{y^{(2)}, y^{(3)}, y^{(4)}\}$.

$P_{\max} = 3.5$, $\Delta T_{\max} = 45$ °C, $h = 8.26$ W/m^2 °C, $T_{pl} = 20$ °C, $T_o = 5$ °C, $T_{amb} = 5$ °C, and $T_{i1} = 10$ °C, $i \in \{1, \ldots, 83\}$. Moreover, the specific heat capacity of air at constant pressure is $C_p = 1.004$ kJ/kgK, the specific heat capacity of air at constant volume is $C_v = 0.717$ kJ/kgK, and the air density is $\rho_{\text{air}} = 1.225$ kg/m^3. The modeling uncertainty associated with each subsystem is modeled as $r^s(t) = 5\% T_{pl} \sin(0.1t)$ °C and $r^{(i)}(t) = 5\% T_{amb} \sin(0.1t)$ °C, $i \in \{1, \ldots, 83\}$. For simulation purposes, the noise corrupting the sensor output is simulated by a uniform random variable with $n^s(t) = [-3\% y^s_{\text{ref}}, 3\% y^s_{\text{ref}}]$ and $n^{(i)}(t) = [-3\% y^{(i)}_{\text{ref}}, 3\% y^{(i)}_{\text{ref}}]$, where y^s_{ref} and $y^{(i)}_{\text{ref}}$ are the set points of temperatures selected as $y^s_{\text{ref}} = 55$ °C and $y^{(i)}_{\text{ref}} = 24$ °C, $i \in \mathbb{N} = \{1, \ldots, 83\}$. The design parameters of the fault diagnosis methodology are selected as follows: $L^s = 5$, $\rho^s = 1$, $\zeta^s = 40$, $L^{(i)} = 5$, for all $i \in \mathbb{N}$, $\rho^{(j)} = 1.1$, $\zeta^{(j)} = 22$, $j \in D = \{i | 5i, i \in \{1, \ldots, 16\}\}$, $\zeta^{(j)} = 15$, $j \in \mathbb{N} \setminus \{D \cup \{81, 82, 83\}\}$ and $L^{(81)} = L^{(82)} = L^{(83)} = 15$, $\rho^{(81)} = \rho^{(82)} = \rho^{(83)} = 1.1$, $\zeta^{(81)} = \zeta^{(82)} = \zeta^{(83)} = 12$. The 83-zone HVAC system is simulated for 4 hours with initial conditions $T_{st}(0) = 30$ °C and $T_{z_i}(0) = 22$ °C, $i \in \{1, \ldots, 83\}$ and a single fault scenario is executed with multiple simultaneous sensor faults such as $f^{(j)}(t_f^{(j)}) = -15\% y^{(j)}_{\text{ref}}$ at $t^{(j)} = 2h$, $j \in \mathcal{J} = \{2, 18, 27, 42, 57, 58, 60, 73, 83\}$, where $\mathcal{J}$ contains the indices of the faulty temperature sensors. The zones with the faulty sensors are indicated with a red square in Fig. 5.

In Fig. 6 the ARRs of the sensor fault diagnosis agents $\mathcal{M}^s$

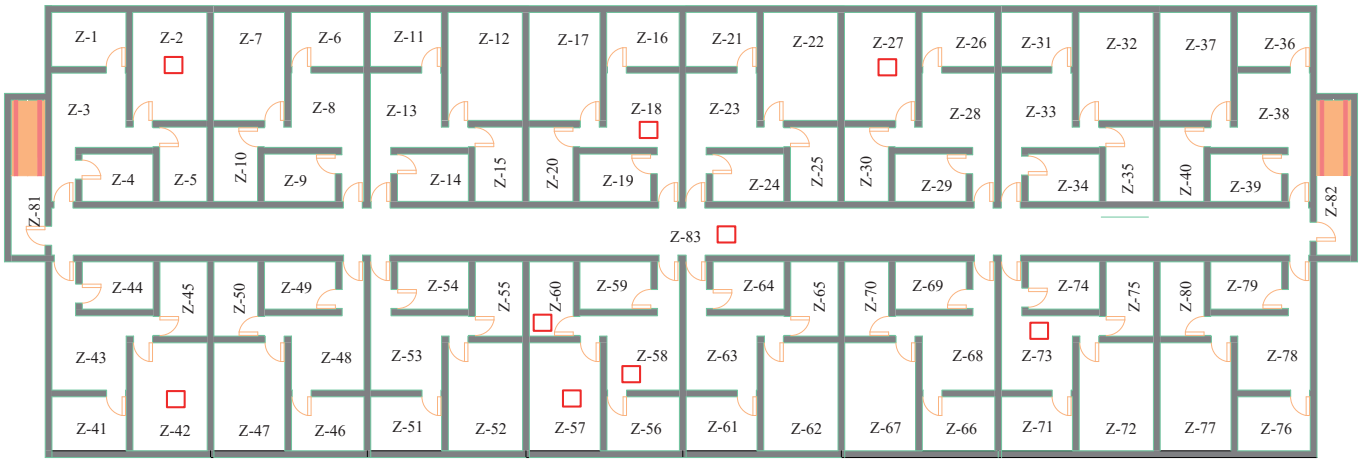


Fig. 5. Down-view of a 83-zone building. Red squared boxes denote the zones with the faulty sensors.

TABLE V
PARAMETERS OF THE 83-ZONE HVAC SYSTEM: ZONES 1 – 5 (1-ST APARTMENT), 81 (LEFT STAIR HALL), AND 83 (CORRIDOR)

Parameter	Value	Unit
$\{C_{st}, C_{z_1}, C_{z_2}, C_{z_3}, C_{z_4}, C_{z_5}, C_{z_{81}}, C_{z_{83}}\}$	{8370, 29.96, 57.71, 54.38, 26.63, 26.63, 3819, 30 557}	kJ/°C
$\{U_{1,\max}, U_{2,\max}, U_{3,\max}, U_{4,\max}, U_{5,\max}, U_{81,\max}, U_{83,\max}\}$	{3700, 7125.9, 6714.8, 3700, 3700, 7400, 59 200}	kg/h
$\{a_{z_1}, a_{z_2}, a_{z_3}, a_{z_4}, a_{z_5}, a_{z_{81}}, a_{z_{83}}\}$	740	kJ/h°C
$\{a_{z_{1,2}}, a_{z_{1,3}}, a_{z_{2,3}}, a_{z_{2,5}}, a_{z_{3,4}}, a_{z_{3,5}}\}$	50	kJ/h°C
$\{A_{w_1}, A_{w_2}, A_{w_3}, A_{w_4}, A_{w_5}, A_{w_{81}}, A_{w_{83}}\}$	{31.21, 43.69, 54.09, 29.72, 29.72, 45.74, 297.24}	m^2
$\{A_{d_{1,3}}, A_{d_{2,3}}, A_{d_{3,4}}, A_{d_{3,5}}, A_{d_{3,81}}, A_{d_{81,83}}\}$	1.951	m^2

and $\mathcal{M}^{(j)}$, $j \in M = \{2, 17, 18, 27, 42, 43, 57, 58, 59, 60, 72, 73, 81, 83\}$ are presented. Note that due to space limitation we have not included the results of all 83 agents. Specifically, each plot of Fig. 6 contains the residuals ε^s, $\varepsilon^{(j)}$, the adaptive thresholds $\bar{\varepsilon}^s$, $\bar{\varepsilon}^{(j)}$ and the decision detection signals D^s, $D^{(j)}$, $j \in M$. Note that sensor fault diagnosis agents $\mathcal{M}^{(2)}$, $\mathcal{M}^{(18)}$, $\mathcal{M}^{(27)}$, $\mathcal{M}^{(42)}$, $\mathcal{M}^{(57)}$, $\mathcal{M}^{(58)}$, $\mathcal{M}^{(60)}$, $\mathcal{M}^{(73)}$ and $\mathcal{M}^{(83)}$ detected the corresponding local sensor faults, while the remainder agents $\mathcal{M}^s$ and $\mathcal{M}^{(j)}$, $j \in \mathbb{N} \setminus \mathcal{J}$ do not detect any sensor fault. From Fig. 6 it can be noticed that the adaptive threshold in (36) is affected by the local sensor faults, while the adaptive thresholds in (37) are affected by both local and neighboring sensor faults.

Every agent that detects sensor fault activates the isolation process (see Section III-B). For example, for the sensor fault isolation process executed by the agent $\mathcal{M}^{(60)}$ the sensor fault signature matrix $F^{(60)}$ is designed and a part of it is presented in Table VI. The observed pattern $\Phi^{(60)}$ at $t = 2.015$ h is

$$\begin{aligned}\Phi^{(60)}(2.015) &= [D^s, D^{(60)}, D^{(55)}, D^{(57)}, D^{(58)}, D^{(59)}, D^{(83)}] \\ &= [0, 1, 0, 1, 1, 0, 1] \end{aligned} \tag{65}$$

and is compared to all theoretical patterns given by the columns of the sensor fault signature matrix $F^{(60)}$ and the agent $\mathcal{M}^{(60)}$ contracts the diagnosis set $\Upsilon^{(60)}$

$$\begin{aligned}\Upsilon^{(60)} = \{&f^{(60)}, f^{(58)}, f^{(60,57)}, f^{(60,78)}, f^{(60,83)}, f^{(57,58)}, f^{(57,83)}, \\ &f^{(58,83)}, f^{(60,57,58)}, f^{(60,57,83)}, f^{(60,58,83)}, f^{(57,58,83)}, \\ &f^{(60,57,58,83)}\}\end{aligned} \tag{66}$$

where $f^{(i,j)}$ represents $f^{(i,j)} = \{f^{(i)}, f^{(j)}\}$. Note that $\mathcal{M}^{(60)}$ can be affected by $2^{|\mathcal{K}_{60}|+2} - 1 = 127$ combinations of sensor faults, however the diagnosis set $\Upsilon^{(60)}$ narrows down the combinations to 13.

Fig. 7 presents the reference points y^s_{ref}, $y^{(j)}_{\text{ref}}$ (black dashed line), the sensor measurements y^s, $y^{(j)}$ (green solid line), the actual temperatures T_{st}, T_{z_j} (red dashed-doted line) and the estimations $\hat{T}_{st}$, $\hat{T}_{z_j}$ (blue dotted line) of the subsystems Σ^s, $\Sigma^{(j)}$, $j \in M$, respectively. It is noted that for those subsystems that the sensor fault occurs locally (e.g., $f^{(1)}$ is the local sensor fault of $\Sigma^{(1)}$) the actual temperature (red dashed-dotted line) deviates from their corresponding reference point (black dashed line). Furthermore, it can be observed that also some zones with healthy local sensor are affected by sensor faults occurring in sensors of neighboring subsystems. For example, the temperature in subsystems $\Sigma^{(17)}$, $\Sigma^{(43)}$, and $\Sigma^{(59)}$ deviate from their corresponding reference point although there is no local sensor fault. This is due to the distributed control scheme that is implemented, where each controller aggregates local and neighboring sensor measurements in order to obtain the local control input, thus the temperature of a zone can be affected also by neighboring sensor faults. Also it is worth mentioning that the corresponding neighboring monitoring agents of the affected subsystems (i.e., $\Sigma^{(17)}$, $\Sigma^{(43)}$, $\Sigma^{(59)}$) as illustrated in Fig. 6 do not detect the sensor faults occurred in their neighboring subsystems (i.e., Σ^s, $\Sigma^{(j)}$, $j \in M$). This is due to the fact that the ARR of each distributed sensor fault diagnosis agent is more sensitive to the occurrence of the local sensor fault and less sensitive to the occurrence of a propagated sensor fault. Further, we can observe that even if

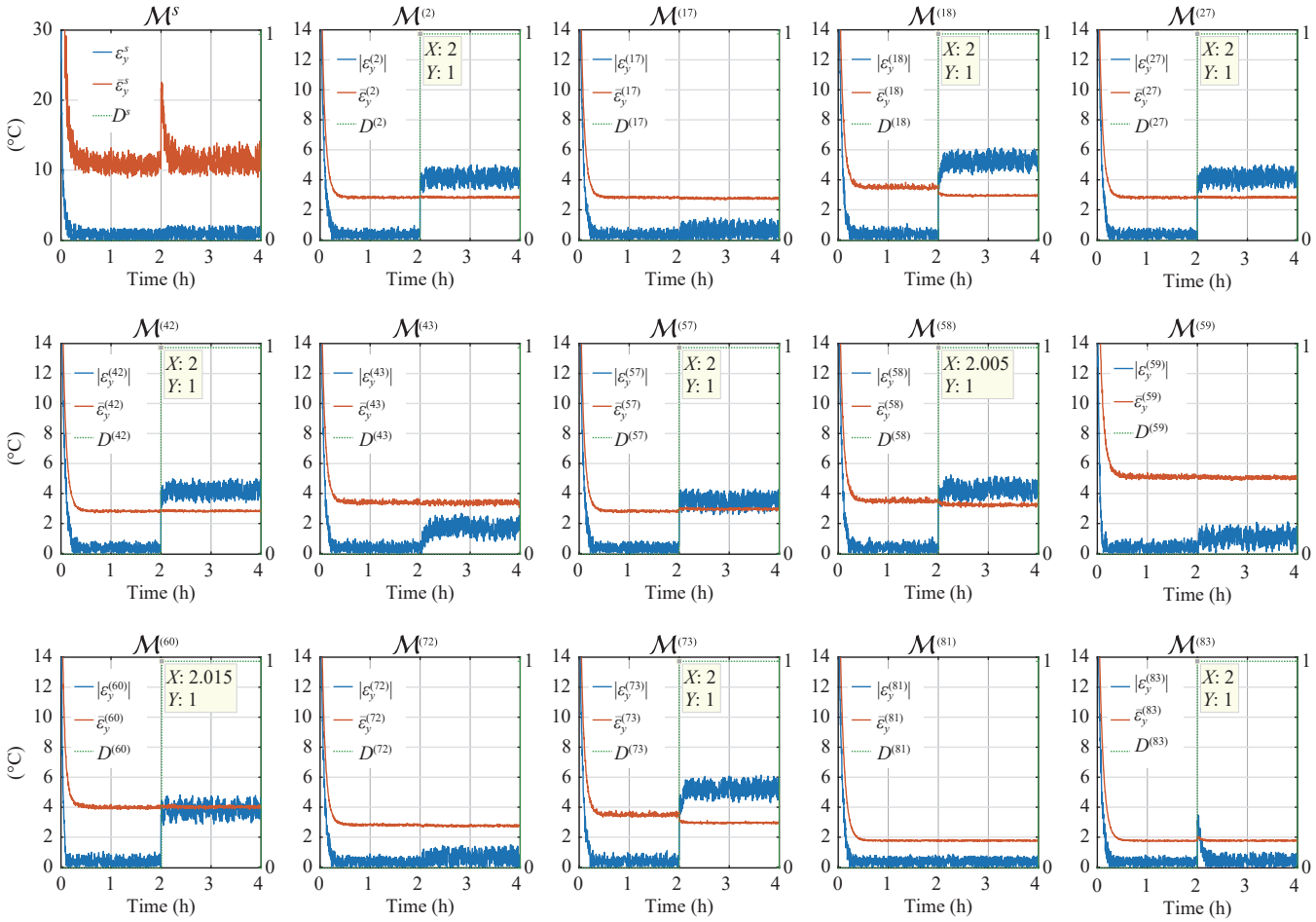


Fig. 6. ARRs of agents $\mathcal{M}^s$ and $\mathcal{M}^{(j)}$, $j \in M$. The residuals ε^s and $\varepsilon^{(j)}$ (blue line), adaptive thresholds $\bar{\varepsilon}^s$ and $\bar{\varepsilon}^{(j)}$ (red line) and boolean decision signals D^s and $D^{(j)}$ (green dotted line) for $j \in M$ are presented.

TABLE VI
THE SENSOR FAULT SIGNATURE MATRIX OF THE AGENT $\mathcal{M}^{(60)}$

	f^s	$f^{(60)}$	$f^{(55)}$	$f^{(57)}$	$f^{(58)}$	$f^{(59)}$	$f^{(83)}$	$f^{(60,57,58,83)}$
$\mathcal{E}^s$	1	*	*	*	*	*	*	*
$\mathcal{E}^{(60)}$	*	1	*	*	*	*	*	1
$\mathcal{E}^{(55)}$	*	*	1	0	0	0	*	*
$\mathcal{E}^{(57)}$	*	*	0	1	*	0	0	1
$\mathcal{E}^{(58)}$	*	*	0	*	1	*	*	1
$\mathcal{E}^{(59)}$	*	*	0	0	*	1	*	*
$\mathcal{E}^{(83)}$	*	*	*	0	*	*	1	1

the actual temperatures of $\Sigma^{(17)}$, $\Sigma^{(43)}$, $\Sigma^{(59)}$ are affected by neighboring faults (i.e., do not track their corresponding reference temperature), both the estimation and measurements of the temperatures are close to the actual temperature. We may infer that the residuals of the neighboring agents are not severely affected from propagated sensor faults, and thus it is more possible to detect a local sensor fault that to detect a sensor fault occurred in a neighboring subsystem. To conclude, the design of the proposed methodology allows to detect and isolate sensor faults even if the use of a distributed control scheme is affected by the propagation a sensor fault.

In order to investigate the effectiveness of the proposed sensor fault diagnosis method, we implemented numerous simulation scenarios modifying the range of noise corrupting the sensor measurements; i.e., $n^{(i)}(t)$ satisfies Assumption 1 with $[0.5\%y^{(i)}_{\text{ref}} \;\; 12\%y^{(i)}_{\text{ref}}]$ for all $i \in \{1,\ldots,83\}$. For the multiple sensor fault scenario denoted with the red squared boxes in Fig. 5, we run 100 times the same simulation while keeping the sensor noise magnitude of all 83 air temperature sensors the same. The simulated sensor faults occur at $t_f^{(j)} = 0.5$ h with $f^{(j)} = 15\%y^{(j)}_{\text{ref}}$ for $j \in \{2,18,27,42,57,58,60,73,83\}$ and the simulation time is 1 h. Fig. 8 shows the percentage of detected local sensor faults (%), given by

$$P_D^{(i)} = \frac{\text{No. of Detected Local Sensor Faults}}{\text{No. of Total Generated Local Sensor Faults}} \times 100\% \tag{67}$$

for each agent with respect to the local sensor noise variance $n^{(i)}(t)$. Specifically, each blue dot in Fig. 8 corresponds to the instances (from the 100 simulations obtained for each sensor noise variance $n^{(i)}(t)$) that the corresponding sensor fault diagnosis agent detected the presence of the local sensor fault. As illustrated, the percentage of detected local sensor faults of the sensor fault diagnosis agents is decreasing as the variance of sensor noise is increasing. Note that detection decision of each agent is not only affected by the noise from its local sensor but it is also affected by sensor noise from its neighbouring sub-

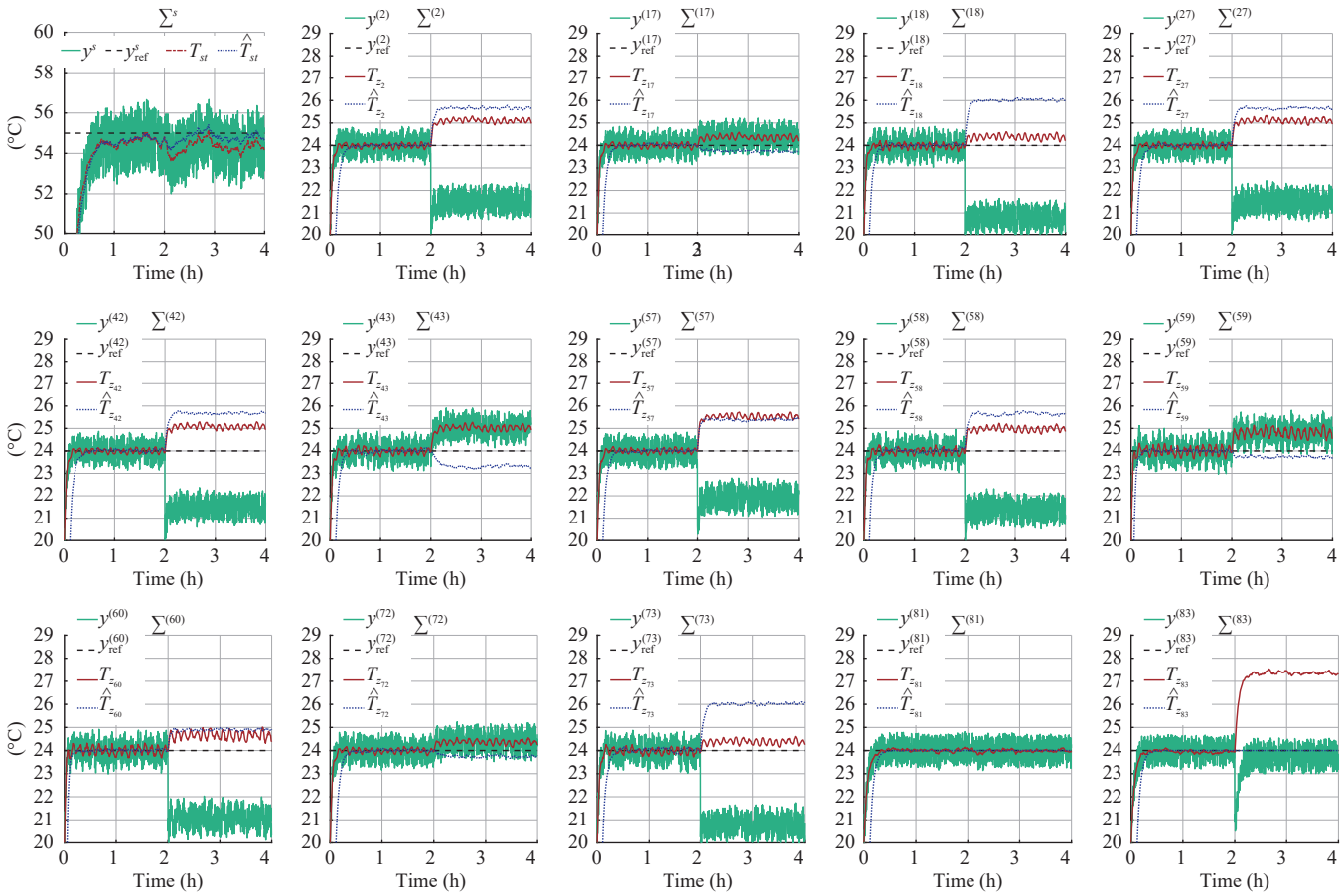


Fig. 7. The temperature reference points y^s_{ref}, $y^{(j)}_{\text{ref}}$ (black dashed line), the sensor measurements y^s, $y^{(j)}$ (green solid line), the temperatures T_{st}, T_{z_j} (red dashed-doted line) and the estimations $\hat{T}_{st}$, $\hat{T}_{z_j}$ (blue dotted line) of subsystems Σ^s and $\Sigma^{(j)}$ for all $j \in M$.

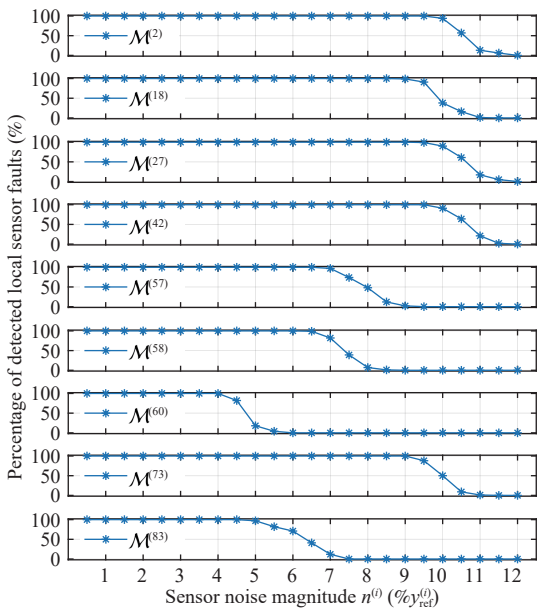


Fig. 8. Percentage of detected local sensor faults with respect to local sensor noise variance $n^{(i)}(t)$. Each blue dot corresponds to the times that the corresponding diagnosis agent detected the presence of the local sensor fault from the 100 simulations obtained for each sensor noise variance $n^{(i)}(t)$. Note that the percentage of sensor noise variance is the same for all sensors in the building.

systems (see (23)–(27) and (37)). Therefore, the agents that monitor zones that have the same number of neighbouring zones (i.e., same $|\mathcal{K}_i|$) and same design properties (see Table V) may have a similar percentage of detected local sensor faults (see $\mathcal{M}^{(2)}$ with $\mathcal{M}^{(42)}$). However, agents that have the same $|\mathcal{K}_i|$ and same design properties, may not have a similar percentage of detected local sensor faults (see $\mathcal{M}^{(18)}$ with $\mathcal{M}^{(58)}$), since due to the distributed topology of the agents, the detection decision can be affected by sensor fault from neighbouring subsystems (i.e., $\mathcal{S}^{(57)}$ and $\mathcal{S}^{(60)}$).

VI. Conclusions

The formulation of large-scale, complex HVAC systems as networks of interconnected subsystems allows the design of scalable distributed model-based sensor fault diagnosis methodologies. The design process of each distributed agent consists of: 1) the sensor fault detection that is based on the generation of ARRs constructed by residuals (resulted by discrepancies of the output and the estimated output of each subsystem) and thresholds that bound the residuals under healthy conditions; and 2) the sensor fault isolation that is obtained using a sensor fault signature matrix which is constructed based on the connectivity of the fault diagnosis agents and allows to eliminate the number of possible locations of the sensor faults. The distributed design of the proposed fault diagnosis method is analyzed in terms of

robustness, detectability, and scalability. The methodology is evaluated under a multiple sensor fault scenario for a large-scale HVAC system consists of 83 building zones. Further, the sensitivity of the proposed method is evaluated with numerous simulation scenarios modifying the sensor noise variance.

It is important to note that the proposed distributed sensor fault diagnosis algorithm can be also applied for diagnosing process or actuator faults. Specifically, the same algorithm is able to detect process and actuator faults, however, the isolation process needs to be modified or extended in order to distinguish between the different types of faults, i.e., process, actuator or sensor faults.

APPENDIX A

The bound $\bar{g}^s$ is designed to bound the difference $\tilde{g}^s$ defined in (18). Given (2), $\tilde{g}^s = \omega(P(T_{st}) - P(y^s))$ where $\omega = U_{st,\max}/C_{st}$ and

$$P(T_{st}) = \begin{cases} (1+\lambda(\kappa - T_{st})), & T_{st}(t) \le \kappa \\ 1, & T_{st} > \kappa \end{cases} \tag{68}$$

$$P(y_H^s) = \begin{cases} \left(1+\lambda\left(\kappa - y_H^s\right)\right), & y_H^s \le \kappa \\ 1, & y_H^s > \kappa \end{cases} \tag{69}$$

with $\lambda = (P_{\max}-1)/\Delta T_{\max}$ and $\kappa = \Delta T_{\max} + T_o$. The variable T_{st} is unknown but belongs to a known interval; i.e., under healthy conditions ($f^s = 0$), (29) is valid, so based on Assumption 2, $T_{st} \in \left[y_H^s - \bar{n}^s, y_H^s + \bar{n}^s\right]$. Due to this inclusion, by applying interval arithmetic we have the following cases

1) if $y_H^s \le \kappa - \bar{n}^s$, then $T_{st} \le \kappa$ and $|\tilde{g}^s| = |\omega\lambda(y_H^s - T_{st})| = |\omega\lambda n^s| \le \omega\lambda\bar{n}^s$;

2) if $y_H^s > \kappa + \bar{n}^s$, then $T_{st}(t) \le \kappa$ and $|\tilde{g}^s| = 0$;

However if $y^s \in (\kappa - \bar{n}^s, \kappa + \bar{n}^s]$, T_{st} may satisfy either $T_{st} > \kappa$ or $T_{st} \le \kappa$. Thus, we need to investigate the following cases:

3) if $y_H^s \in (\kappa, \kappa + \bar{n}^s]$ and $T_{st} \le \kappa$, then $|\tilde{g}^s| = |\omega\lambda(\kappa - T_{st})|$;

4) if $y_H^s \in (\kappa, \kappa + \bar{n}^s]$ and $T_{st} > \kappa$, then $|\tilde{g}^s| = 0$;

5) if $y_H^s \in (\kappa - \bar{n}^s, \kappa]$ and $T_{st} \le \kappa$, then $|\tilde{g}^s| = |\omega\lambda(y_H^s - T_{st})| = |\omega\lambda n^s|$;

6) if $y_H^s \in (\kappa - \bar{n}^s, \kappa]$ and $T_{st} > \kappa$ then $|\tilde{g}^s| = \omega\lambda|\kappa - y_H^s|$.

Given that $T_{st} \in \left[y_H^s - \bar{n}^s, y_H^s + \bar{n}^s\right]$, it yields that $\omega\lambda|\kappa - T_{st}| \in \omega\lambda\left[|\kappa - y_H^s - \bar{n}^s|, |\kappa - y_H^s + \bar{n}^s|\right]$. Therefore, in Cases 3 and 4 $|\tilde{g}^s| \le \max(|\kappa - y_H^s - \bar{n}^s|, |\kappa - y_H^s + \bar{n}^s|)$ irrespective of the exact value of T_{st}. In Case 6, since $y_H^s \in (\kappa - \bar{n}^s, \kappa]$, then $\kappa - y_H^s > 0$ and $\kappa - y_H^s < \bar{n}^s$, we have $|\tilde{g}^s| = \omega\lambda|\kappa - y_H^s| \le \omega\lambda\bar{n}^s$. The following inclusion functions give the bound $\bar{g}^s$. So, in both Cases 5 and 6, $|\tilde{g}^s| \le \omega\lambda\bar{n}^s$. In summary,

$$\bar{g}^s(y_H^s) = \begin{cases} 0, & y_H^s > \kappa + \bar{n}^s \\ \omega\lambda \max\left(|\kappa - y_H^s - \bar{n}^s|, \right. & \\ \left. |\kappa - y_H^s + \bar{n}^s|\right), & y_H^s(t) \in (\kappa, \kappa + \bar{n}^s] \\ \omega\lambda\bar{n}^s, & y_H^s \le \kappa. \end{cases} \tag{70}$$

APPENDIX B

The bound $\bar{\mu}^{(i)}$ is designed to bound the difference $\tilde{\mu}^{(i)} = \mu^{(i)}(T_{z_i}, T_{z_j}) - \mu^{(i)}(y_H^{(i)}, y_H^{(j)})$ where $y_H^{(i)}$ and $y_H^{(j)}$ satisfy (33). Let us define

$$\mu^{(i)}(T_{z_i}(t), T_{z_j}(t)) = \chi_2(t)\chi_1(t) \tag{71}$$

where $\chi_1 = \sqrt{|T_{z_j} - T_{z_i}|}$ and $\chi_2 = \text{sgn}(T_{z_j} - T_{z_i})\max(T_{z_i}, T_{z_j})$. Based on Assumption 2 and (33), $T_{z_i} \in \left[y_H^{(i)} - \bar{n}^{(i)}, y_H^{(i)} + \bar{n}^{(i)}\right]$, for all $i \in \mathbb{N}$. Taking into account the monotonicity of χ_1 and applying interval arithmetic, we obtain

$$\chi_1 \in \left[\underline{\chi}_1^{(i)}, \bar{\chi}_1^{(i)}\right] \tag{72}$$

$$\left[\underline{\chi}_1^{(i)}, \bar{\chi}_1^{(i)}\right] = \begin{cases} \left[\sqrt{\alpha - \beta}, \sqrt{\alpha + \beta}\right], & \text{if } \alpha(t) > \beta \\ \left[\sqrt{|\alpha + \beta|}, \sqrt{|\alpha - \beta|}\right], & \text{if } \alpha(t) < -\beta \\ \left[0, \sqrt{\max(|\alpha - \beta|, |\alpha + \beta|)}\right], & \text{if } |\alpha(t)| \le \beta \end{cases} \tag{73}$$

where $\alpha(t) = y_H^{(j)}(t) - y_H^{(i)}(t)$, $\beta = \bar{n}^{(j)} + \bar{n}^{(i)}$. Following the same procedure, we have

$$\chi_2 \in \left[\underline{\chi}_2^{(i)}, \bar{\chi}_2^{(i)}\right] \tag{74}$$

$$\left[\underline{\chi}_2^{(i)}, \bar{\chi}_2^{(i)}\right] = \begin{cases} \left[y_H^{(j)} - \bar{n}^{(j)}, y_H^{(j)} + \bar{n}^{(j)}\right], & \text{if } \alpha > \beta \\ \left[-y_H^{(i)} - \bar{n}^{(i)}, -y_H^{(i)} + \bar{n}^{(i)}\right], & \text{if } \alpha < -\beta \\ [\min(W), \max(W)], & \text{if } |\alpha| \le \beta \end{cases} \tag{75}$$

$$W = \{-\min(w_1, w_2), -\max(w_1, w_2), \min(w_1, w_2), \max(w_1, w_2)\} \tag{76}$$

where $w_1 = y_H^{(j)}(t) + \bar{n}^{(j)}$ and $w_2 = y_H^{(i)}(t) + \bar{n}^{(i)}$. Based on (71), (72), and (74), it yields

$$\mu^{(i)}(T_{z_i}, T_{z_j}) \in [\underline{\chi}^{(i)}, \bar{\chi}^{(i)}] \tag{77}$$

with

$$\underline{\chi}^{(i)} = \min(\underline{\chi}_1^{(i)}\underline{\chi}_2^{(i)}, \underline{\chi}_1^{(i)}\bar{\chi}_2^{(i)}, \bar{\chi}_1^{(i)}\underline{\chi}_2^{(i)}, \bar{\chi}_1^{(i)}\bar{\chi}_2^{(i)}) \tag{78}$$

$$\bar{\chi}^{(i)} = \max(\underline{\chi}_1^{(i)}\underline{\chi}_2^{(i)}, \underline{\chi}_1^{(i)}\bar{\chi}_2^{(i)}, \bar{\chi}_1^{(i)}\underline{\chi}_2^{(i)}, \bar{\chi}_1^{(i)}\bar{\chi}_2^{(i)}). \tag{79}$$

Using (77) and applying interval arithmetic results in $\tilde{\mu}^{(i)} \in [\underline{\chi}^{(i)} - \upsilon^{(i)}, \bar{\chi}^{(i)} - \upsilon^{(i)})]$ with $\upsilon^{(i)} = \mu^{(i)}(y_H^{(i)}, y_H^{(j)})$. The upper bound that satisfies $|\tilde{\mu}^{(i)}| \le \bar{\mu}^{(i)}(y_H^{(i)}, y_H^{(j)})$ is computed as

$$\bar{\mu}^{(i)}(y_H^{(i)}, y_H^{(j)}) = \max(|\underline{\chi}^{(i)} - \upsilon^{(i)}|, |\bar{\chi}^{(i)} - \upsilon^{(i)}|). \tag{80}$$

REFERENCES

[1] N. E. Klepeis, W. C. Nelson, W. R. Ott, J. P. Robinson, A. M. Tsang, P. Switzer, J. V. Behar, S. C. Hern, and W. H. Engelmann, "The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants," *J. Exposure Analysis and Environmental Epidemiology*, vol. 11, no. 3, pp. 231–252, 2001.

[2] J. Sun and Y. Zhang, "Towards an energy efficient architecture in smart building," in *Proc. Int. Conf. Computational Intelligence and Communication Networks*, 2015, pp. 1589−1592.

[3] A. Capozzoli, F. Lauro, and I. Khan, "Fault detection analysis using data mining techniques for a cluster of smart office buildings," *Expert Systems with Applications*, vol. 42, no. 9, pp. 4324–4338, 2015.

[4] G. Boracchi, M. Michaelides, and M. Roveri, "Detecting contaminants in smart buildings by exploiting temporal and spatial correlation," in *Proc. IEEE Symp. Series Computational Intelligence*, 2015, pp. 601−608.

[5] M. Pătraşcu and M. Drăgoicea, "Integrating agents and services for control and monitoring: managing emergencies in smart buildings," in *Proc. 3rd Int. Workshop Service Orientation in Holonic and Multiagent*

Manufacturing and Robotics, 2013, vol. 544, pp. 209−224.

[6] M. Kintner-Meyer, M. R. Brambley, T. Carlon, and N. Bauman, "Wireless sensors: technology and cost-savings for commercial buildings," *Teaming for Efficiency*: *Proc. the ACEEE Summer Study on Energy Efficiency in Buildings*, vol. 7, no. 8, pp. 121–134, 2002.

[7] R. Isermann, *Fault-Diagnosis Systems: An Introduction From Fault Detection to Fault Tolerance*. Springer Science & Business Media, 2006.

[8] J. Schein, S. T. Bushby, N. S. Castro, and J. M. House, "A rule-based fault detection method for air handling units," *Energy and Buildings*, vol. 38, no. 12, pp. 1485–1492, 2006.

[9] H. Yang, S. Cho, C.-S. Tae, and M. Zaheeruddin, "Sequential rule based algorithms for temperature sensor fault detection in air handling units," *Energy Conversion and Management*, vol. 49, no. 8, pp. 2291–2306, 2008.

[10] Y. Zhao, J. Wen, and S. Wang, "Diagnostic Bayesian networks for diagnosing air handling units faults - part II: faults in coils and sensors," *Applied Thermal Engineering*, vol. 90, pp. 145–157, 2015.

[11] M. Sampath, R. Sengupta, S. Lafortune, and K. Sinnamohideen, "Failure diagnosis using discrete-event models," *IEEE Trans. Control Systems Technology*, vol. 4, no. 2, pp. 105–124, 1996.

[12] S. Katipamula and M. R. Brambley, "Review article: methods for fault detection, diagnostics, and prognostics for building systemsa review, part I," *HVAC&R Research*, vol. 11, no. 1, pp. 3–25, 2005.

[13] S. Wang and F. Xiao, "AHU sensor fault diagnosis using principal component analysis method," *Energy and Buildings*, vol. 36, no. 2, pp. 147–160, 2004.

[14] Z. Du and X. Jin, "Detection and diagnosis for sensor fault in HVAC systems," *Energy Conversion and Management*, vol. 48, no. 3, pp. 693–702, 2007.

[15] M. Kumar and I. N. Kar, "Fault detection and diagnosis of airconditioning systems using residuals," in *Proc. 10th IFAC Int. Symp. Dynamics and Control of Process Systems*, 2013, pp. 607−612.

[16] A. Beghi, L. Cecchinato, L. Corso, M. Rampazzo, and F. Simmini, "Process history-based fault detection and diagnosis for VAVAC systems," in *Proc. IEEE Int. Conf. Control Applications*, 2013, pp. 1165–1170.

[17] J. Liang and R. Du, "Model-based fault detection and diagnosis of HVAC systems using support vector machine method," *Int. Journal of Refrigeration*, vol. 30, no. 6, pp. 1104–1114, 2007.

[18] T. Mulumba, A. Afshari, K. Yan, W. Shen, and L. K. Norford, "Robust model-based fault diagnosis for air handling units," *Energy and Buildings*, vol. 86, pp. 698–707, 2015.

[19] S. Wang and Y. Chen, "Fault-tolerant control for outdoor ventilation air flow rate in buildings based on neural network," *Building and Environment*, vol. 37, no. 7, pp. 691–704, 2002.

[20] Z. Du, X. Jin, and Y. Yang, "Fault diagnosis for temperature, flow rate and pressure sensors in VAV systems using wavelet neural network," *Applied Energy*, vol. 86, no. 9, pp. 1624–1631, 2009.

[21] S. Wang and J.-B. Wang, "Robust sensor fault diagnosis and validation in HVAC systems," *Trans. of Institute of Measurement and Control*, vol. 24, no. 3, pp. 231–262, 2002.

[22] M. Padilla, D. Choinière, and J. A. Candanedo, "A model-based strategy for self-correction of sensor faults in variable air volume air handling units," *Science and Technology for the Built Environment*, vol. 21, no. 7, pp. 1018–1032, 2015.

[23] X. F. Liu and A. Dexter, "Fault-tolerant supervisory control of VAV airconditioning systems," *Energy and Buildings*, vol. 33, no. 4, pp. 379–389, 2001.

[24] C. H. Lo, P. T. Chan, Y. K. Wong, a. B. Rad, and K. L. Cheung, "Fuzzy genetic algorithm for automatic fault detection in HVAC systems," *Applied Soft Computing J.*, vol. 7, no. 2, pp. 554–560, 2007.

[25] H. Yoshida, S. Kumar, and Y. Morita, "Online fault detection and diagnosis in VAV air handling unit by RARX modeling," *Energy and Buildings*, vol. 33, no. 4, pp. 391–401, 2001.

[26] J. C. M. Yiu and S. Wang, "Multiple ARMAX modeling scheme for forecasting air conditioning system performance," *Energy Conversion and Management*, vol. 48, no. 8, pp. 2276–2285, 2007.

[27] W. J. Turner, A. Staino, and B. Basu, "Residential HVAC fault detection using a system identification approach," *Energy and Buildings*, vol. 151, pp. 1–17, 2017.

[28] X. B. Yang, X. Q. Jin, Z. M. Du, Y. H. Zhu, and Y. B. Guo, "A hybrid model-based fault detection strategy for air handling unit sensors," *Energy and Buildings*, vol. 57, pp. 132–143, 2013.

[29] M. Bonvini, M. D. Sohn, J. Granderson, M. Wetter, and M. A. Piette, "Robust on-line fault detection diagnosis for HVAC components based on nonlinear state estimation techniques," *Applied Energy*, vol. 124, pp. 156–166, 2014.

[30] B. T. Thumati, M. A. Feinstein, J. W. Fonda, A. Turnbull, F. J. Weaver, M. E. Calkins, and S. Jagannathan, "An online model-based fault diagnosis scheme for HVAC systems," in *Proc. IEEE Int. Conf. Control Applications*, 2011, pp. 70−75.

[31] P. M. Papadopoulos, V. Reppa, M. M. Polycarpou, and C. G. Panayiotou, "Distributed diagnosis of actuator and sensor faults in HVAC systems," in *Proc. 20th IFAC World Congr.*, 2017, pp. 4293−4293.

[32] H. Shahnazari, P. Mhaskar, J. M. House, and T. I. Salsbury, "Modeling and fault diagnosis design for HVAC systems using recurrent neural networks," *Computers and Chemical Engineering*, vol. 216, pp. 189–203, 2019.

[33] Y. Chen and L. Lan, "Fault detection, diagnosis and data recovery for a real building heating/cooling billing system," *Energy Conversion and Management*, vol. 51, no. 5, pp. 1015–1024, 2010.

[34] R. Yan, Z. J. Ma, G. Kokogiannakis, and Y. Zhao, "A sensor fault detection strategy for air handling units using cluster analysis," *Automation in Construction*, vol. 70, no. 1, pp. 77–88, 2016.

[35] Q. Zhou, S. W. Wang, and Z. J. Ma, "A model-based fault detection and diagnosis strategy for HVAC systems," *Int. J. Energy Research*, vol. 33, no. 10, pp. 903–918, 2009.

[36] S. Wang, Q. Zhou, and F. Xiao, "A system-level fault detection and diagnosis strategy for HVAC systems involving sensor faults," *Energy and Buildings*, vol. 42, no. 4, pp. 477–490, 2010.

[37] D. Sklavounos, E. Zervas, O. Tsakiridis, and J. Stonham, "A subspace identification method for detecting abnormal behavior in HVAC systems," *J. Energy*, vol. 2015, pp. 1−12, 2015.

[38] V. Gunes, S. Peter, and T. Givargis, "Improving energy efficiency and thermal comfort of smart buildings with HVAC systems in the presence of sensor faults," in *Proc. 17th IEEE Int. Conf. High Performance Computing and Communications, 7th Int. Symp. Cyberspace Safety and Security, and 12th Int. Conf. Embedded Software and Systems*, 2015, pp. 945−950.

[39] V. Reppa, P. Papadopoulos, M. M. Polycarpou, and C. G. Panayiotou, "Distributed detection and isolation of sensor faults in HVAC systems," in *Proc. Mediterranean Conf. Control and Automation*, 2013, pp. 401−406.

[40] P. M. Papadopoulos, V. Reppa, M. M. Polycarpou, and C. G. Panayiotou, "Distributed adaptive estimation scheme for isolation of sensor faults in multi-zone HVAC systems," in *Proc. 9th IFAC Symp. Fault Detection, Supervision and Safety for Technical Processes*, 2015, pp. 1146−1151.

[41] V. Reppa, P. Papadopoulos, M. M. Polycarpou, and C. G. Panayiotou, "A distributed architecture for HVAC sensor fault detection and

isolation," *IEEE Trans. Control Systems Technology*, vol. 23, no. 4, pp. 1323–1337, Jul. 2015.

[42] P. M. Papadopoulos, V. Reppa, M. M. Polycarpou, and C. G. Panayiotou, "Distributed adaptive sensor fault tolerant control for smart buildings," in *Proc. 54th IEEE Conf. Decision and Control*, 2015, pp. 3143−3148.

[43] V. Reppa, M. M. Polycarpou, and C. G. Panayiotou, "Sensor Fault Diagnosis," *Foundations and Trends in Systems and Control*, vol. 3, no. 1–2, pp. 1–248, 2016.

[44] S. Riverso, F. Boem, G. Ferrari-Trecate, and T. Parisini, "Fault diagnosis and control-reconfiguration in large-scale systems : a plug-andplay approach," in *Proc. IEEE Conf. Decision and Control*, 2014, pp. 4977−4982.

[45] V. Reppa, P. Papadopoulos, M. M. Polycarpou, and C. G. Panayiotou, "A distributed virtual sensor scheme for smart buildings based on adaptive approximation," in *Proc. IEEE Int. Joint Conf. Neural Networks*, 2014, pp. 99−106.

[46] M. Zaheer-Uddin, "Temperature control of multizone indoor spaces based on forecast and actual loads," *Building and Environment*, vol. 29, no. 4, pp. 485–493, 1994.

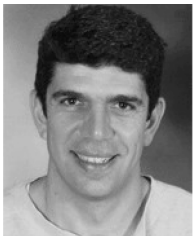
[47] M. Zaheer-Uddin and R. V. Patel, "Optimal tracking control of multizone indoor environmental spaces," *J. Dynamic Systems, Measurement, and Control*, vol. 117, no. 3, pp. 292–303, 1995.

[48] E. Witrant, S. Mocanu, O. Sename, and Others, "A hybrid model and MIMO control for intelligent buildings temperature regulation over WSN," in *Proc. 8th IFAC Workshop Time Delay Systems*, 2009, pp. 420−425.

Panayiotis M. Papadopoulos (S'13) received the B.Sc. and M.Sc. degrees from the University of Cyprus, Cyprus, in 2012 and 2014, respectively, where he is currently pursuing the Ph.D. degree, all in electrical engineering. His current research interests include fault diagnosis for distributed systems, nonlinear control theory, intelligent systems, and adaptive fault-tolerant control in HVAC Systems. He is a Student Member of ASHRAE and Chair of IEEE University of Cyprus Student Branch from Feb. 2015 to Feb. 2018. He has been awarded an Honorable Mention as one of the finalist papers for the Young Author Award presented at the *10th IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes, SAFEPROCESS 2018*.

Vasso Reppa (M'06) has been an Assistant Professor in the Department of Maritime and Transport Technology of Delft University of Technology, The Netherlands since 2018. She obtained the doctorate (2010) in electrical and computer engineering from the University of Patras, Greece. In 2009 she joined IBM Zurich Research Laboratory in Switzerland as a student intern. From 2011 to 2017, she was a Research Associate (now Research Affiliate) with the KIOS Research and Innovation Center of Excellence in Cyprus. In 2013, Dr. Reppa was awarded the Marie Curie Intra European Fellowship and worked as a Research Fellow in CentraleSupelec of the University Paris-Saclay, France from 2014 to 2016. She was a Visiting Researcher at Imperial College, the UK and at University of Newcastle, Australia in 2015 and 2016, respectively. Her research interests include distributed fault diagnosis and fault tolerant control, adaptive learning, observer-based estimation, and applications of autonomous systems in transport, smart buildings, and robotics. Dr Reppa has led the implementation of work packages of several research and development projects (e.g., FP7, INTERREG, H2020, NWO).

Marios M. Polycarpou (F'06) is a Professor of electrical and computer engineering and the Director of the KIOS Research and Innovation Center of Excellence at the University of Cyprus, Cyprus. He received the B.A. degree in computer science and the B.Sc. in electrical engineering, both from Rice University, USA in 1987, and the M.S. and Ph.D. degrees in electrical engineering from the University of Southern California, USA, in 1989 and 1992, respectively. His teaching and research interests include intelligent systems and networks, adaptive and learning control systems, computational intelligence, fault diagnosis, and critical infrastructure systems. Dr. Polycarpou has published more than 350 articles in refereed journals, edited books, and refereed conference proceedings, and co-authored 7 books. He is also the Holder of 6 patents. Prof. Polycarpou is a Fellow of IEEE and IFAC. He was the Recipient of the 2016 IEEE Neural Networks Pioneer Award and the 2014 Best Paper Award for the journal *Building and Environment* (Elsevier). Prof. Polycarpou served as the President of the IEEE Computational Intelligence Society (2012–2013), the President of the European Control Association (2017–2019), and the Editor-in-Chief of the *IEEE Transactions on Neural Networks and Learning Systems* (2004–2010). He has participated in more than 70 research projects/grants, funded by several agencies and industry in Europe and the United States, including the prestigious European Research Council (ERC) Advanced Grant and the EU Teaming project. Prof. Polycarpou is an elected Founding Member of the Cyprus Academy of Sciences, Letters, and Arts.



Christos G. Panayiotou (SM'10) is a Professor with the Electrical and Computer Engineering (ECE) Department at the University of Cyprus, Cyprus. He is also the Deputy Director of the KIOS Research and Innovation Center of Excellence for which he is also a Founding Member. Christos has received the B.Sc. and a Ph.D. degrees in electrical and computer engineering from the University of Massachusetts, USA, in 1994 and 1999, respectively. He also received the MBA from the Isenberg School of Management, at the aforementioned university in 1999. Before joining the University of Cyprus in 2002, he was a Research Associate at the Center for Information and System Engineering (CISE) and the Manufacturing Engineering Department at Boston University, USA (1999–2002). His research interests include modeling, control, optimization and performance evaluation of discrete event and hybrid systems, cyber-physical systems, event detection and localization, fault diagnosis, smart camera networks, wireless, Ad-hoc and sensor networks, resource allocation, intelligent transportation networks, and intelligent buildings. Christos has published more than 230 papers in international refereed journals and conferences and is the Recipient of the 2014 Best Paper Award for the journal *Building and Environment* (Elsevier). He is an Associate Editor for the *IEEE Transactions of Intelligent Transportation Systems*, the *IEEE Transactions on Control Systems Technology*, the Conference Editorial Board of the *IEEE Control Systems Society*, the *Journal of Discrete Event Dynamical Systems,* and the *European Journal of Control*. He held several positions in organizing committees and technical program committees of numerous international conferences. He has also served as Chair of various subcommittees of the Education Committee of the IEEE Computational Intelligence Society.