

# A Distributed Virtual Sensor Scheme for Smart Buildings based on Adaptive Approximation

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

**Abstract**—This paper presents the design of a methodology for diagnosing sensor faults in heating, ventilation and air-conditioning (HVAC) systems, and compensating their effects on the distributed control architecture. The proposed methodology is developed in a distributed framework, considering a multi-zone HVAC system as a set of interconnected, nonlinear subsystems. For each of the interconnected subsystems, we design a local virtual sensor agent that can detect and isolate faults in its monitored sensors and provide sensor fault estimations for correcting the faulty measurements. Adaptive estimation schemes are implemented in each local virtual sensor agent, using adaptive approximation models for learning the unknown fault function. Simulation results are used for illustrating the effectiveness of the proposed methodology applied to a two-zone HVAC system.

## I. INTRODUCTION

SMART BUILDINGS incorporate embedded intelligence based on computer technology, aiming at autonomously adapting the evolving building environment in order to increase energy efficiency and cost effectiveness, improve comfort, productivity and safety, and enhance system robustness and reliability [1]. The heating ventilation and air-conditioning (HVAC) system constitutes a basic component of buildings, necessary for providing a high quality and healthy environment for the occupants. However, the HVAC system may increase the energy waste in commercial buildings when it is improperly controlled and its performance is degraded, especially due to the occurrence of faults. In that case, the energy waste was estimated to be 15% to 30%, according to the Electrical and Mechanical Services Department of Hong Kong [2].

A feasible and reliable solution for managing the energy consumption and, at the same time, maintaining comfort conditions in buildings is the implementation of fault diagnosis and accommodation mechanisms for detecting and isolating faults in HVAC systems and compensating their effects on the feedback control scheme [3]. During the last two decades, various methodologies have been developed for active fault tolerant control (AFTC) in HVAC systems. Most of these methodologies have focused on faults in actuators and the plant of the HVAC system. However, diagnosing sensor faults and ensuring the fault tolerant performance of the HVAC system are becoming key challenging problems, due to the

large number of sensors used for monitoring and control of energy consumption and living conditions in smart buildings.

An efficient approach for ensuring the proper operation of HVAC systems under sensor fault conditions is the employment of AFTC schemes based on virtual sensors. The design of virtual sensors relies on developing mathematical models of the process implemented in software, which are used to reconstruct, correct or predict the faulty or missing measurements [4]. In HVAC systems, which are typically highly complex, spatially distributed and with a large number of interconnected components, the utilization of model-based virtual sensors provides a more appealing approach compared to physical redundancy. Using physical sensors approach implies additional cost for installation and maintenance, while, if they are not added to the system during the initial HVAC installation, especially in the electromechanical part, invasive actions by specialized personnel are required. This may delay the initialization of the proper HVAC operation after the fault isolation or even risk a serious damage of the electromechanical part when accessing the system.

Several researchers have investigated the design of virtual sensor schemes, combining information from healthy sensors with static or dynamic analytical models, such as observers, aiming at reconstructing the output of a faulty sensor or correcting the faulty output using the estimation of the sensor fault [5]–[8]. In HVAC systems, there is a significant research activity in designing virtual sensors following a data-driven modeling approach with the goal of predicting the output of a faulty sensor [9]–[11]. However, these methods require a large amount of data collected under various normal operating conditions for synthesizing the virtual sensor models. An alternative approach to the design of virtual sensors for HVAC systems relies on the use of static models based on first-principles (see [4] and the references therein). This type of virtual sensors is more appropriate for monitoring or fault identification than for feedback control. There are very few virtual sensor schemes based on correcting the faulty output using the sensor fault estimation [12], while to the authors' best knowledge, no work has yet been developed on distributed virtual sensor schemes.

The objective and main contribution of this work is the development of a distributed virtual sensor scheme for compensating the effects of sensor faults on the distributed control of HVAC system using an adaptive approximation approach. In previous work [13], the authors have developed a distributed methodology for detecting and isolating sensor faults in a multi-zone HVAC system, based on local sensor fault diagnosis agents, where every agent is used to

The authors are with the KIOS Research Center for Intelligent Systems and Networks, Department of Electrical & Computer Engineering, University of Cyprus, Nicosia, 1678, Cyprus (email: {vreppa,ppapad01,mpolycar,christosp}@ucy.ac.cy).

This work is funded by the European Research Council Advanced Grant FAULT-ADAPTIVE (ERC-AdG-291508).

monitor one of the HVAC subsystems. In this work, the designed agents aim not only to detect and isolate sensor faults, but also to compensate their effects by generating virtual signals of the faulty sensors. To this end, we design local virtual sensor (LVS) agents using dynamic estimation schemes with real-time learning capabilities. The rationale behind the learning approach is the on-line approximation of the unmodeled nonlinear sensor faults by using adaptive approximation models (e.g. sigmoidal neural networks and radial basis functions) [14]–[16].

The development of a distributed virtual sensor scheme can significantly contribute in preventing the propagation of sensor faults in a distributed control architecture for HVAC systems. Moreover, a distributed virtual sensor methodology can provide a more suitable approach for HVAC systems in smart buildings with multiple sensors compared to a centralized approach, since it may reduce the computational complexity of the AFTC algorithms, increase their reliability with respect to security threats and be scalable in case of system expansion.

This paper is organized as follows. The HVAC description and problem formulation is described in Section II. The architecture and the design details of the proposed distributed virtual sensor approach are presented in Section III. Simulation results of the application of the proposed virtual sensor architecture to a two-zone HVAC system are provided in Section IV, followed by concluding remarks in Section V.

## II. HVAC DESCRIPTION AND PROBLEM FORMULATION

Consider a HVAC system, which consists of  $N$  separated zones (e.g. dormitory rooms, classrooms) and the electromechanical part. The basic components of the electromechanical part of the HVAC, shown in Fig. 1 are the cooling coil, the chiller and the chilled water tank, the fan, the supply and return ducts and the variable air volume (VAV) boxes. The cooling coil is connected to the chiller through the chiller water tank, which regulates the water inserted to the cooling coil.

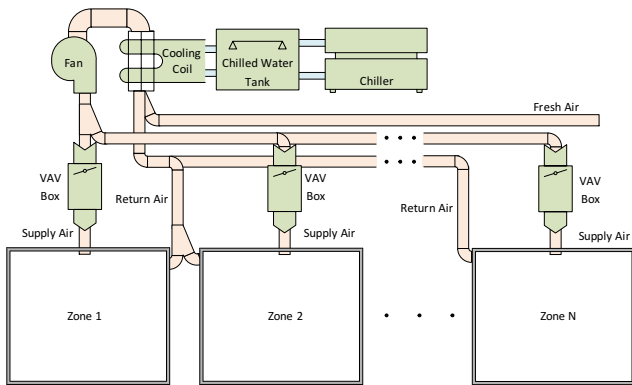


Fig. 1. Schematic of a  $N$ -zone HVAC system.

The temperature dynamics in each zone, cooling coil and chiller water tank can be modeled based on the fundamental

mass and energy conservation equations [17], [18]. The  $N$ -zone HVAC system can be regarded as a set of  $N + 1$  interconnected, nonlinear subsystems that correspond to the electromechanical part, comprised of the cooling coil and chiller water tank, and the  $N$  building zones. Let  $T_{z_I}$  ( $^{\circ}\text{C}$ ) be the temperature of the  $I$ -th zone,  $I \in \{1, \dots, N\}$ ,  $T_1^e$  ( $^{\circ}\text{C}$ ) be the output air temperature from cooling coil and  $T_2^e$  ( $^{\circ}\text{C}$ ) be the temperature of the water in the chiller storage tank, while the variable  $Q_{a_I}$  ( $\text{m}^3/\text{sec}$ ) is the volumetric flow rate of air entering into the  $I$ -th zone and  $\chi$  ( $\text{m}^3/\text{sec}$ ) is the chilled water mass flow rate. The dynamics of the subsystem that corresponds to the electromechanical part, denoted by  $\Sigma^e$ , can be expressed as:

$$\Sigma^e : \quad \frac{dT^e(t)}{dt} = A^e T^e(t) + \gamma^e(\chi(t)) + h^e(T^e(t), T_z(t), Q_a(t)), \quad (1)$$

where  $T^e = [T_1^e, T_2^e]^T$ ,  $T_z = [T_{z_1}, \dots, T_{z_N}]^T$ ,  $Q_a = [Q_{a_1}, \dots, Q_{a_N}]^T$  and

$$A^e = \begin{bmatrix} -\frac{U_{cc} A_{cc}}{M_{cc} C_v} & \frac{Q_w \rho_w C_{pw}}{M_{cc} C_v} \\ 0 & -\frac{Q_w \rho_w C_{pw} + U_t A_t}{V_t \rho_w C_{pw}} \end{bmatrix} \quad (2)$$

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

$$h^e(T_1^e, T_z, Q_a) = \begin{bmatrix} h_1^e(T_1^e, T_z, Q_a) \\ 0 \end{bmatrix} \quad (4)$$

$$h_1^e(T_1^e, T_z, Q_a) = \left( \frac{\rho_a C_{pa}}{M_{cc} C_v} \sum_{I=1}^N Q_{a_I} - \frac{U_{cc} A_{cc}}{M_{cc} C_v} \right) \frac{1}{N} \sum_{I=1}^N T_{z_I} + \frac{\rho_a}{M_{cc} C_v} (h_{fg} - C_{pa})(w_z - w_{ao}) \sum_{I=1}^N Q_{a_I} - \frac{\rho_a C_{pa}}{M_{cc} C_v} T_1^e \sum_{I=1}^N Q_{a_I}. \quad (5)$$

It is noted that the first two terms of (1) represent the local dynamics of  $\Sigma^e$ , while  $h^e$  characterizes the interconnection dynamics between  $\Sigma^e$  and  $\{\Sigma^{(1)}, \dots, \Sigma^{(N)}\}$ , where  $\Sigma^{(I)}$  corresponds to the dynamics of the  $I$ -th zone for all  $I \in \{1, \dots, N\}$  expressed as:

$$\Sigma^{(I)} : \quad \frac{dT_{z_I}(t)}{dt} = A^{(I)} T_{z_I}(t) + \gamma^{(I)}(T_{z_I}(t), Q_{a_I}(t)) + h^{(I)}(T_{z_I}(t), Q_{a_I}(t)) + \eta^{(I)}(t), \quad (6)$$

where  $A^{(I)} = -\frac{U_{z_I} A_{z_I}}{M_{z_I} C_v}$ ,  $\eta^{(I)} = \frac{1}{M_{z_I} C_v} \tilde{T}_{z_I}$  and

$$\gamma^{(I)}(T_{z_I}, Q_{a_I}) = -\frac{\rho_a C_{pa}}{M_{z_I} C_v} T_{z_I} Q_{a_I} + \frac{U_{z_I} A_{z_I}}{M_{z_I} C_v} T_{amb}, \quad (7)$$

$$h^{(I)}(T_1^e, Q_{a_I}) = \frac{\rho_a C_{pa}}{M_{z_I} C_v} T_1^e Q_{a_I}. \quad (8)$$

Again, the first two terms of (6),  $A^{(I)} T_{z_I}$  and  $\gamma^{(I)}(T_{z_I}, Q_{a_I})$  correspond to the local dynamics of  $\Sigma^{(I)}$ , while  $h^{(I)}$  represents the interconnection dynamics between  $\Sigma^{(I)}$  and  $\Sigma^e$ . The value  $\tilde{T}_{z_I}(t)$  ( $^{\circ}\text{C}/\text{sec}$ ) represents the rate of internal heat

change, due to occupants and appliances from the  $I$ -th zone. In this work, it is assumed that the ambient temperature  $T_{amb}$  ( $^{\circ}\text{C}$ ) is known. For a more detailed description of the HVAC system please refer to [13], [17].

The control inputs to the  $N$ -zone HVAC system are the volumetric flow rate of air to each zone  $Q_{a_i}$  and the chilled water mass flow rate to the storage tank  $\chi$ , generated by feedback controllers based on some reference signals  $r^{(I)}$  and  $r^e$  respectively. By controlling these inputs, the objective is to achieve the desired temperature in each building zone (for occupants' comfort) and in the cooling coil (for energy efficiency). In each of the  $N$  zones, there exists a sensor measuring the zone temperature  $T_{z_i}$ , while two sensors are available in the electromechanical part of the HVAC, measuring the temperature of the air exiting the cooling coil  $T_1^e$  and the temperature of the chilled water in the tank  $T_2^e$ . Specifically, the  $I$ -th subsystem  $\Sigma^{(I)}$ ,  $I \in \{1, \dots, N\}$ , is monitored and controlled using a temperature sensor, denoted by  $\mathcal{S}^{(I)}$ , characterized by the output  $y^{(I)} \in \mathbb{R}$ ; i.e.,

$$\mathcal{S}^{(I)} : \quad y^{(I)}(t) = T_{z_i}(t) + d^{(I)}(t) + f^{(I)}(t), \quad (9)$$

where  $d^{(I)} \in \mathbb{R}$  denotes the noise corrupting the measurements  $y^{(I)}$  of sensor  $\mathcal{S}^{(I)}$  and  $f^{(I)} \in \mathbb{R}$  represents the possible abrupt sensor fault, described by:

$$f^{(I)}(t) = \beta(t - t_f^{(I)})\phi^{(I)}(Q_{a_i}(t)), \quad (10)$$

where  $\beta(t)$  is the time profile and  $\phi^{(I)}(\cdot)$  is the (unknown) function of the sensor fault that occurs at the (unknown) time instant  $t_f^{(I)}$ . The time profile of the abrupt fault is modeled as  $\beta(t) = 0$  for  $t < 0$  and  $\beta(t) = 1$  for  $t \geq 0$ .

The nonlinear subsystem  $\Sigma^e$  is monitored and controlled using a sensor set  $\mathcal{S}^e$  that includes two temperature sensors  $\mathcal{S}^e\{1\}$  and  $\mathcal{S}^e\{2\}$ , characterized by

$$\mathcal{S}^e\{1\} : \quad y_1^e(t) = T_1^e(t) + d_1^e(t) + f_1^e(t) \quad (11)$$

$$\mathcal{S}^e\{2\} : \quad y_2^e(t) = T_2^e(t) + d_2^e(t) + f_2^e(t), \quad (12)$$

where  $y_j^e \in \mathbb{R}$  is the sensor output,  $d_j^e \in \mathbb{R}$  denotes the noise corrupting the measurements of sensor  $\mathcal{S}^e\{j\}$  and  $f_j^e \in \mathbb{R}$  represents the possible abrupt sensor fault described by:

$$f_j^e(t) = \beta(t - t_{f_j}^e)\phi_j^e(\chi(t)), \quad j = 1, 2 \quad (13)$$

where  $\phi_j^e(\cdot)$  is the (unknown) function of the sensor fault that occurs at the (unknown) time instant  $t_{f_j}^e$ .

The objective of this work is to design a methodology using an adaptive approximation approach for detecting and isolating sensor faults that may occur in one of the subsystems and compensate their effects on the distributed control of the  $N$ -zone HVAC system.

### III. DISTRIBUTED VIRTUAL SENSOR SCHEME

The design of the proposed distributed virtual sensor scheme is realized as follows. Taking into account the  $N + 1$  subsystems, defined by (1)-(8), the first step is to design a local virtual sensor (LVS) agent for each of the interconnected subsystems; i.e. the agent  $\mathcal{M}^e$  dedicated to subsystem  $\Sigma^e$  and the agent  $\mathcal{M}^{(I)}$  dedicated to subsystem

$\Sigma^{(I)}$ ,  $I \in \{1, \dots, N\}$ , as shown in Fig. 2. Each LVS agent has access to the control input and sensor output data of the underlying subsystem, while it may exchange information with some other agents. The agents  $\mathcal{M}^{(1)}, \dots, \mathcal{M}^{(N)}$  monitoring the building zones do not exchange any information. The exchanged information is associated with the form of the physical (state) and input interconnections and is data-driven. In addition, event-driven communication is established between the agents  $\mathcal{M}^e$  and each  $\mathcal{M}^{(I)}$  based on the distributed sensor fault diagnosis decision logic executed by each agent, which will be described later on.

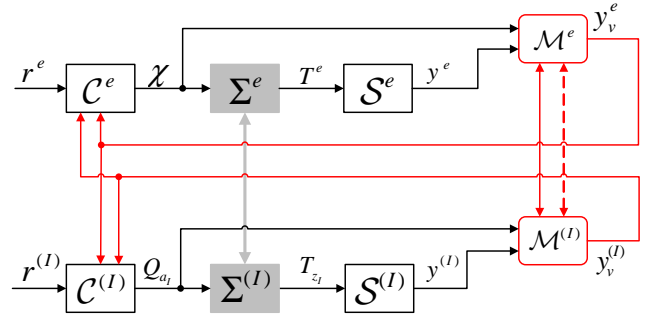


Fig. 2. Distributed virtual sensor scheme for HVAC system, where  $y_v^e$  and  $y_v^{(I)}$  correspond to virtual sensor signals.

Assuming the occurrence of sensor faults in one of the subsystems, the first goal of the proposed distributed virtual sensor scheme is the detection and isolation of the faults, given the bidirectional exchange of sensor information between each agent  $\mathcal{M}^{(I)}$  and  $\mathcal{M}^e$ . Considering the occurrence of multiple sensor faults in the sensor set of the electromechanical part, two modules are designed in the agent  $\mathcal{M}^e$  such that the  $j$ -th module, denoted by  $\mathcal{M}_j^e$  is dedicated to the sensor  $\mathcal{S}^e\{j\}$ ,  $j = 1, 2$ . When at least one agent  $\mathcal{M}^{(I)}$  or  $\mathcal{M}^e$  detects sensor faults, all agents  $\mathcal{M}^{(I)}$  transmit their decisions to  $\mathcal{M}^e$ . Due to the propagation of sensor faults through the exchange of information, the distributed sensor fault isolation procedure is conducted by processing combinatorially the decisions of the agents monitoring the building zones and the agent monitoring the electromechanical part. The second goal of the proposed distributed virtual sensor scheme is to compensate the effects of sensor faults on the control of the HVAC system by using on-line approximation models of the unknown sensor faults affecting one of the subsystems.

In order to accomplish these tasks, the agent  $\mathcal{M}^{(I)}$  and the modules  $\mathcal{M}_1^e$  and  $\mathcal{M}_2^e$  are designed using an adaptive nonlinear estimation scheme. Based on the generated estimation model and using local input and sensor output data, as well as transmitted information, each agent  $\mathcal{M}^{(I)}$ , (correspondingly modules  $\mathcal{M}_1^e$  and  $\mathcal{M}_2^e$ ) generates residuals and adaptive thresholds that are further used to formulate analytical redundancy relations (ARRs). The distributed sensor fault detection decision is obtained when the generated ARR are not satisfied. The adaptive approximation of the

isolated sensor faults is carried out by expressing the fault functions  $\phi^{(I)}$  and  $\phi_j^e$  as linearly parametrized approximators (e.g. radial basis functions) [19], [20]; i.e.

$$\phi^{(I)}(Q_{a_I}(t)) = G^{(I)}(Q_{a_I}(t))\theta^{(I)} + \delta_\phi^{(I)}(Q_{a_I}(t)) \quad (14)$$

$$\phi_j^e(\chi(t)) = G_j^e(\chi(t))\theta_j^e + \delta_{\phi_j}^e(\chi(t)) \quad (15)$$

where  $G^{(I)} \in \mathbb{R}^{1 \times q_I}$  and  $G_j^e \in \mathbb{R}^{1 \times q_{e_j}}$  are the regressors of the adaptive approximators and  $\theta^{(I)} \in \mathbb{R}^{q_I}$ ,  $\theta_j^e \in \mathbb{R}^{q_{e_j}}$  are the (unknown) optimal parameter (weight) vectors that minimize the discrepancy between the corresponding unknown function and the adaptive approximator. The residual errors  $\delta_\phi^{(I)}$  and  $\delta_{\phi_j}^e$  are referred to as *minimum functional approximation errors* over the compact sets  $\mathcal{Q}_I$ ,  $\mathcal{X}_I$ , within which  $Q_{a_I}$  and  $\chi$  reside respectively. The on-line approximation models of the sensor faults are used to correct the output of the faulty sensors and the corrected measurements are provided to the distributed control architecture as virtual signals  $y_v^{(I)}$ ,  $y_v^e$  (Fig. 2).

#### A. Adaptive Nonlinear Estimators for Residual Generation

The residual generated by the module  $\mathcal{M}_1^e$ , denoted by  $\varepsilon_{y_1}^e \in \mathbb{R}$ , is defined as

$$\varepsilon_{y_1}^e(t) = y_{v_1}^e(t) - C_1^e \hat{T}^e(t), \quad (16)$$

where  $\hat{T}^e \in \mathbb{R}^2$  is the estimation of  $T^e$  and  $y_{v_1}^e(t)$  is the virtual signal generated by the module  $\mathcal{M}_1^e$ , defined as

$$y_{v_1}^e(t) = y_1^e(t) - G_1^e(\chi(t))\hat{\theta}_1^e(t), \quad (17)$$

with  $\hat{\theta}_1^e(t) \in \mathbb{R}^{q_{e_1}}$  be a vector that contains the adjustable parameters of the adaptive approximation model of  $\phi_1^e$ . The state and parameter estimates  $\hat{T}^e$ ,  $\hat{\theta}_1^e$  are determined based on the adaptive nonlinear estimation model of the module  $\mathcal{M}_1^e$ , selected as follows:

$$\begin{aligned} \dot{\hat{T}}^e(t) &= A^e \hat{T}^e(t) + \gamma^e(\chi(t)) + h^e(y_{v_1}^e(t), y_v(t), Q_a(t)) \\ &\quad + L_1^e \varepsilon_{y_1}^e(t) + \Omega_1^e(t) \hat{\theta}_1^e(t) \end{aligned} \quad (18)$$

$$\begin{aligned} \dot{\Omega}_1^e(t) &= A_{L_1}^e \Omega_1^e(t) + \left[ \frac{\rho_a C_{pa}}{M_{cc} C_v} \sum_{I=1}^N Q_{a_I}(t) \right] G_1^e(\chi(t)) \\ &\quad - L_1^e G_1^e(\chi(t)) \end{aligned} \quad (19)$$

$$\dot{\hat{\theta}}_1^e(t) = \mathcal{P}_1^e \left\{ \Gamma_1^e (C_1^e \Omega_1^e(t) + G_1^e(\chi(t)))^\top \mathcal{D}_1^e [\varepsilon_{y_1}^e(t)] \right\}, \quad (20)$$

where  $L_1^e \in \mathbb{R}^{2 \times 1}$  is the estimator gain matrix, chosen such that  $A_{L_1}^e = A^e - L_1^e C_1^e$  is stable,  $C_1^e = [1, 0]$  and  $y_v(t) = [y_v^{(1)}(t), \dots, y_v^{(N)}(t)]^\top$ , where  $y_v^{(I)}(t)$  denotes the virtual signal transmitted by the agent  $\mathcal{M}^{(I)}$ ,  $I \in \{1, \dots, N\}$ , defined later on. The matrix  $\Omega_1^e(t) \in \mathbb{R}^{2 \times q_{e_1}}$  is a filter, used to ensure the stability property of the adaptive nonlinear estimation scheme [14]. The initial conditions are  $\hat{T}_1^e(0) = 0$ ,  $\hat{\theta}_1^e(0) = 0$  and  $\Omega_1^e(0) = 0$ . In the adaptive law (20),  $\Gamma_1^e \in \mathbb{R}^{q_{e_1} \times q_{e_1}}$  is a symmetric, positive definite learning rate matrix, while the projection operator  $\mathcal{P}_1^e$  restricts the adjustable parameter vector  $\hat{\theta}_1^e(t)$  to a predefined convex

set  $\Theta_1^e$  for preventing parameter drift [19]. The dead-zone operator  $\mathcal{D}_1^e[\cdot]$  is defined as

$$\mathcal{D}_1^e [\varepsilon_{y_1}^e(t)] = \begin{cases} 0, & \text{if } D_1^e(t) = 0 \text{ and } f_1^e \notin \mathcal{I}(t) \\ \varepsilon_{y_1}^e(t), & \text{otherwise} \end{cases} \quad (21)$$

where  $D_1^e(t)$ ,  $\mathcal{I}(t)$  are determined later on in Section III-C.

The residual generated by the module  $\mathcal{M}_2^e$ , denoted by  $\varepsilon_{y_2}^e \in \mathbb{R}$ , is expressed as:

$$\varepsilon_{y_2}^e(t) = y_{v_2}^e(t) - \hat{T}_2^e(t), \quad (22)$$

where  $\hat{T}_2^e \in \mathbb{R}$  is the estimation of  $T_2^e$  and  $y_{v_2}^e(t)$  is the virtual signal generated by the module  $\mathcal{M}_2^e$ , defined as

$$y_{v_2}^e(t) = y_2^e(t) - G_2^e(\chi(t))\hat{\theta}_2^e(t), \quad (23)$$

with  $\hat{\theta}_2^e(t) \in \mathbb{R}^{q_{e_2}}$  defined as a vector that contains the adjustable parameters of the adaptive approximator of  $\phi_2^e$ . The estimates  $\hat{T}_2^e$  and  $\hat{\theta}_2^e$  are computed based on the following adaptive nonlinear estimator:

$$\dot{\hat{T}}_2^e(t) = A_{22}^e \hat{T}_2^e(t) + \gamma_2^e(\chi(t)) + L_2^e \varepsilon_{y_2}^e(t) + \Omega_2^e(t) \hat{\theta}_2^e(t) \quad (24)$$

$$\dot{\Omega}_2^e(t) = A_{L_2}^e \Omega_2^e(t) - L_2^e G_2^e(\chi(t)) \quad (25)$$

$$\dot{\hat{\theta}}_2^e(t) = \mathcal{P}_2^e \left\{ \Gamma_2^e (\Omega_2^e(t) + G_2^e(\chi(t)))^\top \mathcal{D}_2^e [\varepsilon_{y_2}^e(t)] \right\}, \quad (26)$$

where  $A_{22}^e$  is the element  $\{2, 2\}$  of the matrix  $A^e$  given in (2),  $\gamma_2^e$  is the second element of  $\gamma^e$  and  $L_2^e \in \mathbb{R}$  is the estimator gain chosen such  $A_{L_2}^e = A_{22}^e - L_2^e$  is stable. The initial conditions are  $\hat{T}_2^e(0) = 0$ ,  $\hat{\theta}_2^e(0) = 0$  and  $\Omega_2^e(0) = 0$ . The terms  $\Omega_2^e(t) \in \mathbb{R}^{1 \times q_{e_2}}$ ,  $\Gamma_2^e \in \mathbb{R}^{q_{e_2} \times q_{e_2}}$  and  $\mathcal{P}_2^e$  are defined as the terms  $\Omega_1^e$ ,  $\Gamma_1^e$  and  $\mathcal{P}_1^e$  respectively. The dead-zone operator  $\mathcal{D}_2^e[\cdot]$  is defined as

$$\mathcal{D}_2^e [\varepsilon_{y_2}^e(t)] = \begin{cases} 0, & \text{if } D_2^e(t) = 0 \\ \varepsilon_{y_2}^e(t), & \text{otherwise} \end{cases} \quad (27)$$

where  $D_2^e(t)$  is determined later on in Section III-C.

The residual generated by the agent  $\mathcal{M}^{(I)}$ ,  $I \in \{1, \dots, N\}$ , is denoted by  $\varepsilon_y^{(I)} \in \mathbb{R}$  and is described by

$$\varepsilon_y^{(I)}(t) = y_v^{(I)}(t) - \hat{T}_{z_I}(t), \quad (28)$$

where  $\hat{T}_{z_I} \in \mathbb{R}$  is the estimation of  $T_{z_I}$ ,  $I \in \{1, \dots, N\}$ , and  $y_v^{(I)}(t)$  is the virtual signal generated by the module  $\mathcal{M}^{(I)}$ :

$$y_v^{(I)}(t) = y^{(I)}(t) - G^{(I)}(Q_{a_I}(t))\hat{\theta}^{(I)}(t), \quad (29)$$

with  $\hat{\theta}^{(I)}(t) \in \mathbb{R}^{q_I}$  be a vector that contains the adjustable parameters of the adaptive approximator  $\hat{\phi}^{(I)}$ . The estimates  $\hat{T}^{(I)}$  and  $\hat{\theta}^{(I)}$  are generated based on the adaptive nonlinear estimation scheme, structured as

$$\begin{aligned} \dot{\hat{T}}_{z_I}(t) &= A^{(I)} \hat{T}_{z_I}(t) + \gamma^{(I)}(y_v^{(I)}(t), Q_{a_I}(t)) + \Omega^{(I)}(t) \hat{\theta}^{(I)}(t) \\ &\quad + h^{(I)}(y_{v_1}^e(t), Q_{a_I}(t)) + L^{(I)} \varepsilon_y^{(I)}(t) \end{aligned} \quad (30)$$

$$\begin{aligned} \dot{\Omega}^{(I)}(t) &= A_L^{(I)} \Omega^{(I)}(t) + \frac{\rho_a C_{pa}}{M_{cc} C_v} Q_{a_I}(t) G^{(I)}(Q_{a_I}(t)) \\ &\quad - L^{(I)} G^{(I)}(Q_{a_I}(t)) \end{aligned} \quad (31)$$

$$\dot{\hat{\theta}}^{(I)}(t) = \mathcal{P}^{(I)} \left\{ \Gamma^{(I)} \left( \Omega^{(I)}(t) + G^{(I)}(Q_{a_I}(t)) \right)^\top \cdot \mathcal{P}^{(I)} \left[ \varepsilon_y^{(I)}(t) \right] \right\} \quad (32)$$

where  $L^{(I)} \in \mathbb{R}$  is the estimator gain, chosen such that  $A_L^{(I)} = A^{(I)} - L^{(I)}$  is stable, while the terms  $\Omega^{(I)}(t) \in \mathbb{R}^{1 \times q_I}$ ,  $\Gamma^{(I)} \in \mathbb{R}^{q^{(I)} \times q^{(I)}}$  and  $\mathcal{P}^{(I)}$  are defined as the terms  $\Omega_1^e$ ,  $\Gamma_1^e$  and  $\mathcal{P}_1^e$ . The initial conditions are  $\hat{T}^{(I)}(0) = 0$  and  $\hat{\theta}^{(I)}(0) = 0$  and  $\Omega^{(I)}(0) = 0$ . The dead-zone operator  $\mathcal{P}^{(I)}[\cdot]$  is again defined as

$$\mathcal{P}^{(I)} \left[ \varepsilon_y^{(I)}(t) \right] = \begin{cases} 0, & \text{if } D^{(I)}(t) = 0 \text{ and } f^{(I)} \notin \mathcal{I}(t) \\ \varepsilon_y^{(I)}(t), & \text{otherwise} \end{cases} \quad (33)$$

where  $D^{(I)}(t)$ ,  $\mathcal{I}(t)$  are determined later on in Section III-C.

### B. Computation of Adaptive Thresholds

The computation of adaptive thresholds, designed to ensure the robustness of the agents  $\mathcal{M}^e$  and  $\mathcal{M}^{(I)}$  with respect to modeling uncertainty (disturbances, noise, etc), is realized by taking into account the following assumption:

*Assumption 1:* The modeling uncertainty of  $\Sigma^{(I)}$ ,  $I \in \{1, \dots, N\}$  and the measurement noise of each sensor  $\mathcal{S}^{(I)}$  and  $\mathcal{S}^e\{j\}$ ,  $j = 1, 2$  are unknown but uniformly bounded; i.e.  $|\eta^{(I)}(t)| \leq \bar{\eta}^{(I)}$ ,  $|d^{(I)}(t)| \leq \bar{d}^{(I)}$  and  $|d_j^e(t)| \leq \bar{d}_j^e$ , where  $\bar{\eta}^{(I)}$ ,  $\bar{d}_j^{(I)}$ ,  $\bar{d}_j^e$  are known constant bounds.

The bound for the modeling uncertainty is commonly used for distinguishing between disturbances and faults [21], while the bound for the noise terms corresponds to a practical representation of the available knowledge for the noise that is typically provided in a given operation range by sensor manufacturers.

The adaptive threshold implemented in the module  $\mathcal{M}_{y_j}^e$ , denoted by  $\bar{\varepsilon}_{y_j}^e(t)$ ,  $j = 1, 2$ , is computed such that

$$|\varepsilon_{y_j}^e(t)| \leq \bar{\varepsilon}_{y_j}^e(t), \quad (34)$$

where  $\varepsilon_{y_j}^e(t)$  is the residual under healthy conditions, defined through (16)-(20) for  $j = 1$  and (22)-(26) for  $j = 2$  with  $\hat{\theta}_j^e = 0$ ,  $j = 1, 2$ , using the solutions of (18) and (24). Taking into account Assumption 1 and that there exists a known bound  $\bar{T}^e$  such that  $|T^e(0)| \leq \bar{T}^e$ , and positive constants  $\rho_1^e$ ,  $\xi_1^e$  such that  $|C_1^e e^{A_{L_1}^e t}| \leq \rho_1^e e^{-\xi_1^e t}$  for all  $t$ , the adaptive threshold is expressed as:

$$\bar{\varepsilon}_{y_1}^e(t) = \rho_1^e e^{-\xi_1^e t} \bar{T}^e + \bar{d}_1^e + \int_0^t \rho_1^e e^{-\xi_1^e(t-\tau)} \left( |L_1^e| \bar{d}_1^e + \bar{h}^e(\tau) \right) d\tau, \quad (35)$$

where  $\bar{h}^e(t)$  is computed such that  $|h^e(T_1^e(t), T_z(t), Q_a(t)) - h^e(y_1^e(t), y(t), Q_a(t))| \leq \bar{h}^e(t)$ , where  $y = [y^{(1)}, \dots, y^{(N)}]^\top$ ; i.e.,

$$\bar{h}^e(t) = \left| \frac{\rho_a C_{pa}}{M_{cc} C_v} \sum_{I=1}^N Q_{a_I}(t) - \frac{U_{cc} A_{cc}}{M_{cc} C_v} \right| \frac{1}{N} \sum_{I=1}^N \bar{d}^{(I)} + \frac{\rho_a C_{pa}}{M_{cc} C_v} \bar{d}_1^e \left| \sum_{I=1}^N Q_{a_I}(t) \right|. \quad (36)$$

Taking into account (34), the adaptive threshold  $\bar{\varepsilon}_{y_2}^e$  generated by the module  $\mathcal{M}_{y_2}^e$  is described by

$$\bar{\varepsilon}_{y_2}^e(t) = \rho_2^e e^{-\xi_2^e t} \bar{T}_2^e + \bar{d}_2^e + \int_0^t \rho_2^e e^{-\xi_2^e(t-\tau)} |L_2^e| \bar{d}_2^e d\tau, \quad (37)$$

where  $\bar{T}_2^e$  is a known bound such that  $|T_2^e(0)| \leq \bar{T}_2^e$ , and  $\rho_2^e$ ,  $\xi_2^e$  are positive constants such that  $|e^{A_{L_2}^e t}| \leq \rho_2^e e^{-\xi_2^e t}$  for all  $t$ .

The adaptive threshold implemented in the agent  $\mathcal{M}^{(I)}$ , denoted by  $\bar{\varepsilon}_y^{(I)}(t)$ ,  $I \in \{1, \dots, N\}$ , is computed such that

$$|\varepsilon_y^{(I)}(t)| \leq \bar{\varepsilon}_y^{(I)}(t), \quad (38)$$

where  $\varepsilon_y^{(I)}(t)$  is the residual under healthy conditions, defined through (28)-(32) with  $\hat{\theta}^{(I)} = 0$ , as well as the solution of (30). The adaptive threshold  $\bar{\varepsilon}_y^{(I)}(t)$  is described by:

$$\bar{\varepsilon}_y^{(I)}(t) = \rho^{(I)} e^{-\xi^{(I)} t} \bar{T}_{z_I} + \bar{d}^{(I)} + \int_0^t \rho^{(I)} e^{-\xi^{(I)}(t-\tau)} \left( \bar{\eta}^{(I)} + |L^{(I)}| \bar{d}^{(I)} + \frac{\rho_a C_{pa}}{M_{z_I} C_v} (\bar{d}^{(I)} + \bar{d}_1^e) |Q_{a_I}(\tau)| \right) d\tau \quad (39)$$

where  $\bar{T}_{z_I}$  is a known bound such that  $|T_{z_I}(0)| \leq \bar{T}_{z_I}$ ,  $\rho^{(I)}$ ,  $\xi^{(I)}$  are positive constants such that  $|e^{A_{L^{(I)}} t}| \leq \rho^{(I)} e^{-\xi^{(I)} t}$  for all  $t$ .

It is noted that the adaptive thresholds defined in (35), (37) and (39) can be implemented using linear filtering techniques [20].

### C. Distributed Sensor Fault Detection and Isolation Decision Logic

The generated residuals and adaptive thresholds are used to formulate analytical redundancy relations (ARRs), based on which the distributed sensor fault detection and isolation decision logic is designed [22], [23]. Under healthy conditions of the HVAC system, the computed ARRr are always satisfied.

#### 1) Distributed Sensor Fault Detection Decision Logic:

The decision on the occurrence of sensor faults (detection) is obtained by the modules  $\mathcal{M}_1^e$  and  $\mathcal{M}_2^e$ , dedicated to the electromechanical part, and each agent  $\mathcal{M}^{(I)}$ , dedicated to the building zone  $I$ , based on some analytical redundancy relations (ARRs). Specifically, the ARRr of the modules  $\mathcal{M}_1^e$  and  $\mathcal{M}_2^e$  are defined as:

$$\mathcal{E}_j^e : \left| \varepsilon_{y_j}^e(t) \right| - \bar{\varepsilon}_{y_j}^e(t) \leq 0, \quad j = 1, 2 \quad (40)$$

where  $\varepsilon_{y_1}^e$ ,  $\varepsilon_{y_2}^e$  and  $\bar{\varepsilon}_{y_1}^e$ ,  $\bar{\varepsilon}_{y_2}^e$  are defined in (16), (22) and (35), (37), respectively. The module  $\mathcal{M}_{y_j}^e$  infers the presence of sensor faults, when  $\mathcal{E}_j^e$  defined in (40) is violated. The decision of the module  $\mathcal{M}_{y_j}^e$ ,  $j = 1, 2$  is denoted by  $D_j^e(t)$ , described by the following boolean function

$$D_j^e(t) = \begin{cases} 0, & \text{if } t < t_{D_j}^e \\ 1, & \text{if } t \geq t_{D_j}^e \end{cases} \quad (41)$$

where  $t_{D_j}^e$  is the time instant of detection, defined as

$$t_{D_j}^e = \min_t \{ t : |\varepsilon_{y_j}^e(t)| - \bar{\varepsilon}_{y_j}^e(t) > 0 \}. \quad (42)$$

Assuming the sensor fault occurrence in one of the interconnected subsystems, when  $D_1^e(t) = 1$ , the module  $\mathcal{M}_1^e$  detects the occurrence of sensor faults affecting either the sensor  $\mathcal{S}^e\{1\}$  or the sensor  $\mathcal{S}^{(I)}$ ,  $I \in \{1, \dots, N\}$ . When  $D_2^e(t) = 1$ , the module  $\mathcal{M}_2^e$  detects the occurrence of sensor fault affecting  $\mathcal{S}^e\{2\}$ .

The sensor fault detection decision logic of the agent  $\mathcal{M}^{(I)}$ ,  $I \in \{1, \dots, N\}$  is based on the following ARR

$$\mathcal{E}^{(I)} : \left| \varepsilon_y^{(I)}(t) \right| - \bar{\varepsilon}_y^{(I)}(t) \leq 0, \quad I \in \{1, \dots, N\}, \quad (43)$$

where  $\varepsilon_y^{(I)}$  and  $\bar{\varepsilon}_y^{(I)}$  are defined in (28) and (39), respectively. The decision of  $\mathcal{M}^{(I)}$  on the presence of sensor faults  $f^{(I)}$  or  $f_1^e$  is represented by a boolean function, defined as

$$D^{(I)}(t) = \begin{cases} 0, & \text{if } t < t_D^{(I)} \\ 1, & \text{if } t \geq t_D^{(I)} \end{cases} \quad (44)$$

$$t_D^{(I)} = \min_t \{t : |\varepsilon_y^{(I)}(t)| - \bar{\varepsilon}_y^{(I)}(t) > 0\} \quad (45)$$

where  $t_D^{(I)}$  is the time of detection for agent  $\mathcal{M}^{(I)}$ . When  $D^{(I)}(t) = 1$  the agent  $\mathcal{M}^{(I)}$ ,  $I \in \{1, \dots, N\}$  infers that either  $f_1^e$  or  $f^{(I)}$  has occurred.

### 2) Distributed Sensor Fault Isolation Decision Logic:

After fault detection, the sensor fault isolation procedure is initiated. The agent  $\mathcal{M}^e$  isolates sensor faults in the sensor set  $\mathcal{S}^e$  by comparing the observed pattern of sensor faults, defined as  $D^e(t) = [D_1^e(t), D_2^e(t)]^\top$ , to the columns of the sensor fault signature matrix  $F^e$ , presented in Table I. The rows of  $F^e$  correspond to the ARRs  $\mathcal{E}_1^e$  and  $\mathcal{E}_2^e$ , while the columns correspond to the four possible combinations of sensor faults, i.e.  $\mathcal{F}_1^e = \{f_1^e\}$ ,  $\mathcal{F}_2^e = \{f_2^e\}$ ,  $\mathcal{F}_3^e = \{f_1^e, f_2^e\}$  and  $\mathcal{F}_4^e = \{f_z\}$ , where  $f_z$  represents the sensor faults propagated to the agent  $\mathcal{M}^e$  due to the exchange of sensor information with the agents  $\mathcal{M}^{(1)}, \dots, \mathcal{M}^{(N)}$  (i.e.  $f_z = [f^{(1)}, \dots, f^{(N)}]^\top$ ). The  $j$ -th theoretical pattern of the matrix  $F^e$  is defined as  $F_j^e = [F_{1j}^e, F_{2j}^e]^\top$ ,  $j = 1, 2, 3, 4$ , where  $F_{qj}^e = 1$  if at least one sensor fault of the combination  $\mathcal{F}_j^e$  is involved in the ARR  $\mathcal{E}_q^e$ , and  $F_{qj}^e = 0$  otherwise. For isolating

TABLE I  
SENSOR FAULT SIGNATURE MATRIX  $F^e$ .

	$\mathcal{F}_1^e$	$\mathcal{F}_2^e$	$\mathcal{F}_3^e$	$\mathcal{F}_4^e$
$\mathcal{E}_1^e$	1	0	1	1
$\mathcal{E}_2^e$	0	1	1	0

sensor faults, the agent  $\mathcal{M}^e$  checks the consistency between the observed pattern  $D^e(t)$  and the theoretical pattern  $F^e$ . As long as  $D^e(t) = [0, 0]^\top$ , no consistency check is realized; otherwise, the result of the consistency test is the diagnosed sensor fault combination.

In order for the agent  $\mathcal{M}^e$  to distinguish between a sensor fault in the electromechanical part or a sensor fault propagated from an agent  $\mathcal{M}^{(I)}$ , the agent  $\mathcal{M}^e$  requests  $\mathcal{M}^{(I)}$  for all  $I$  to transmit its decision. This event-driven communication is also carried out, when an agent  $\mathcal{M}^{(I)}$  detects sensor faults. The distributed sensor fault isolation procedure executed by the agent  $\mathcal{M}^e$  is based on comparing the

observed pattern  $D(t) = [D_1(t), D_2(t), \dots, D_{N+1}(t)]^\top = [D_1^e(t), D^{(1)}(t), \dots, D^{(N)}(t)]^\top$  to the columns of the sensor fault signature matrix  $F$ , which is designed assuming the occurrence of sensor faults in one of the subsystems. The matrix  $F$  has  $N + 1$  rows that correspond to the ARRs  $\mathcal{E}_1^e, \mathcal{E}^{(1)}, \dots, \mathcal{E}^{(N)}$ , and  $N + 1$  columns that correspond to the  $N + 1$  sensor faults, i.e.  $f_1^e, f^{(1)}, \dots, f^{(N)}$ . The  $j$ -th theoretical pattern of the matrix  $F$  is defined as  $F_j = [F_{1j}, \dots, F_{(N+1)j}]^\top$ ,  $j = 1, \dots, N + 1$ , where  $F_{qj} = 1$  if the sensor fault is involved in the ARR of the  $q$ -th column and affects the local sensor,  $F_{qj} = *$  if the sensor fault is involved in the ARR, but is propagated from another agent, and  $F_{qj} = 0$  if the sensor fault is not involved in the ARR. Using the notations  $F_{qj} = 1$  and  $F_{qj} = *$ , we discriminate the sensitivity of the ARRs to local and propagated sensor faults [23]. The notation  $F_{qj} = *$  indicates that a sensor fault involved in an ARR and propagated from a neighboring agent can explain why an ARR is violated at some time instant, but ARR may happen to be satisfied while this sensor fault has occurred. For example, if  $N = 2$ , the matrix  $F$  is designed as in Table II.

TABLE II  
SENSOR FAULT SIGNATURE MATRIX  $F$ .

	$f_1^e$	$f^{(1)}$	$f^{(2)}$
$\mathcal{E}_1^e$	1	*	*
$\mathcal{E}^{(1)}$	*	1	0
$\mathcal{E}^{(2)}$	*	0	1

For distributed sensor fault isolation, the observed pattern  $D(t)$  is compared to the columns of  $F$ . The result of the consistency test is the determination of the isolation set  $\mathcal{I}$  that contains the possible diagnosed sensor fault combinations. It is noted that  $F_{qj} = *$  is consistent with either  $D_q(t) = 0$  or  $D_q(t) = 1$ . For example, based on Table II if at some time instant  $t$ ,  $D(t) = [1, 0, 0]^\top$ , the isolation set is  $\mathcal{I} = \{f_1^e\}$ . The isolation set may contain more than one possible diagnosed sensor faults. For example, if  $D(t) = [1, 0, 1]^\top$ , the isolation set is  $\mathcal{I}(t) = \{\{f_1^e\}, \{f^{(2)}\}\}$ , inferring that either  $f_1^e$  or  $f^{(2)}$  has occurred. During the consistency test, there may be an observed pattern that is not consistent with the columns of the matrix  $F$ . In this case, the isolation set is determined from the support of the violated ARRs [22], [24].

The resultant isolation set is transmitted to each agent  $\mathcal{M}^{(I)}$  for initiating or not the learning procedure, as described by the dead-zone operator  $\mathcal{D}^{(I)}$  in (33). The set  $\mathcal{I}$  is also exploited in the dead-zone operator of  $\mathcal{D}_1^e$  of the module  $\mathcal{M}_1^e$ , while the dead-zone operator of the module  $\mathcal{M}_2^e$  is activated based on the decision function  $D_1^e(t)$ , defined in (41), which implies the isolation of sensor fault  $f_2^e$ .

## IV. SIMULATION EXAMPLE

The objective of the simulation example is to illustrate the effectiveness of the proposed distributed virtual sensor scheme applied to the HVAC system described in Section II consisting of two zones ( $N = 2$ ). The operation of the



HVAC system is simulated based on equations (1) and (6), using the parameter values provided in [13]. In addition, we consider that the ambient temperature  $T_{amb}$  is periodically varying within 24 hours as  $T_{amb} = 30 + 5 \sin(0.07\pi \cdot t)$ . Two feedback linearization controllers were implemented for the tracking control of the temperature of each zone, aiming at keeping the temperature at 24°C. A backstepping controller was applied for maintaining the temperature of the output air of the cooling coil at 10°C. Following the design of the distributed virtual sensor methodology provided in Section III, we design four estimators, structured as in (30), (18) and (24) with estimator gains:  $L^{(l)} = 3$ ,  $I = 1, 2$ ,  $L_1^e = [5.43, 12.31]^T$  and  $L_2^e = 3$ . The adaptive thresholds given in (35), (37) and (39) are designed using the following parameters:  $\rho^{(l)} = 1$ ,  $\xi^{(l)} = 3$ ,  $I = 1, 2$ ,  $\rho_1^e = 0.7$ ,  $\xi_1^e = 3.86$ ,  $\rho_2^e = 1$  and  $\xi_2^e = 3$ . In this example, it is assumed that an abrupt fault affects the temperature sensor in zone one at the time instant  $t_f^{(1)} = 6$  h, while the change in the sensor output is described by the function  $\phi^{(1)}(t) = 3 - \cos(600 \cdot Q_{a_1}(t))$ .

The distributed SFDI decision logic is applied as described in Section III-C; at the time instant  $t = 6$ h,  $D^{(1)}(t) = 1$  so the agent  $\mathcal{M}^{(1)}$  detects the presence of sensor faults, i.e. infers that there is a fault affecting the temperature sensor in either zone one or the cooling coil. On the contrary, the agent  $\mathcal{M}^{(2)}$  and the modules  $\mathcal{M}_1^e$ ,  $\mathcal{M}_2^e$  do not detect any fault, i.e.  $D^{(2)}(t) = D_1^e(t) = D_2^e(t) = 0$  for all  $t$ . Once the fault is detected, the agent  $\mathcal{M}^{(1)}$  transmits its decision to the agent  $\mathcal{M}^{(e)}$  aiming at isolating the sensor fault. Based on the sensor fault signature matrix  $F$  given in Table II and the generated observed pattern  $D(t) = [D_1^e(t), D^{(1)}(t), D^{(2)}(t)]^T = [0, 1, 0]^T$  for  $t \geq 6$ h, the resultant isolation set is  $\mathcal{I}(t) = \{f^{(1)}\}$ . Qualitatively, we may infer that the effects of sensor fault  $f^{(1)}$  in the decision of the module  $\mathcal{M}_1^{(e)}$  were weak, so they were not detectable by the module  $\mathcal{M}_1^{(e)}$ , which does not detect any fault in contrast to the agent  $\mathcal{M}^{(1)}$ .

The isolation set is transmitted to the agent  $\mathcal{M}^{(1)}$ , which, based on the dead-zone operator defined in (33), initiates the adaptive approximation procedure of the isolated sensor fault  $f^{(1)}$ . The result of the adaptive approximation is shown in Fig. 3, where the approximation model is derived, using a network of 12 radial basis functions (RBF), with fixed centers evenly distributed in the interval  $[0, 0.005]$ , while the width of each RBF is 0.0004. The approximation model of sensor fault  $\hat{\phi}^{(1)}(t) = G^{(1)}(Q_{a_1}(t))\theta^{(1)}(t)$  is used for generating the virtual signal  $y_v^{(1)}(t)$  as defined in (29). The effectiveness of the proposed distributed virtual sensor scheme is illustrated in Fig. 4. Specifically, Fig. 4 presents the actual temperature  $T_{z_1}(t)$  and air-flow rate  $Q_{a_1}(t)$  (black, dashed line), which are respectively compared to the temperature set point and the air-flow rate generated under healthy conditions (green, solid line) with (subfigures 4(a) and 4(c)) and without (subfigures 4(b) and 4(d)) using the virtual sensor signals.

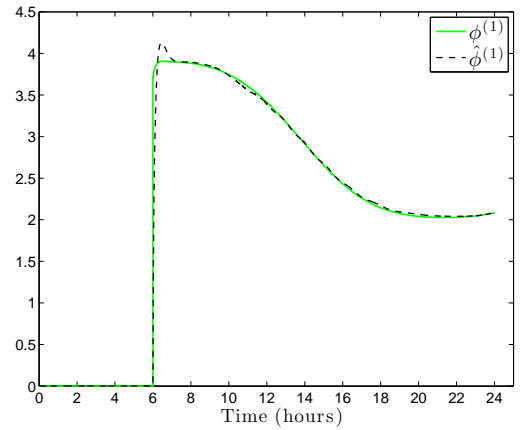


Fig. 3. Fault function  $\phi^{(1)}(t)$  and its approximation model  $\hat{\phi}^{(1)}(t)$ .

## V. CONCLUSIONS

In this paper, we presented a model-based, distributed architecture for compensating the effects of sensor faults in a multi-zone HVAC system using a bank of local virtual sensor (LVS) agents. The first goal of each LVS agent is to detect sensor fault in the local subsystem, while it may exchange information with other agents. The distributed isolation of sensor faults is conducted by combining the decisions of the LVS agents and applying a reasoning-based decision logic. The second goal of each LVS agent is to learn the isolated sensor faults and correct the outputs of the local faulty sensors, providing a virtual signal to the distributed control scheme in order to prevent the sensor fault propagation. Simulation results illustrated the effectiveness of the proposed architecture, by showing the fault tolerant operation of a two-zone HVAC system under sensor fault conditions. Future work will evolve the exhaustive investigation of various sensor fault scenarios for validating the efficiency of the proposed scheme, as well as the extension of the proposed technique for compensating the effects of multiple sensor faults affecting more than one HVAC subsystems.

## REFERENCES

- [1] J. E. Braun, "Intelligent building systems-Past, present, and future," in *Proceedings of American Control Conference*, 2007, pp. 4374–4381.
- [2] Energy Efficiency Office, "Hong Kong energy end-use data," Electrical and Mechanical Services Department (EMSD), The Government of the Hong Kong Special Administrative Region, Tech. Rep., 1999.
- [3] M. Blanke, M. Kinnaert, J. Lunze, and M. Staroswiecki, *Diagnosis and Fault-Tolerant Control*. Springer Verlag, 2003.
- [4] H. Li, D. Yu, and J. E. Braun, "A review of virtual sensing technology and application in building systems," *HVAC&R Research*, vol. 17, no. 5, pp. 619–645, 2011.
- [5] R. Isermann, *Fault-Diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance*. Springer Verlag, 2006.
- [6] N. E. Wu, S. Thavamani, Y. Zhang, and M. Blanke, "Sensor fault masking of a ship propulsion system," *Control Engineering Practice*, vol. 14, no. 11, pp. 1337–1345, 2006.
- [7] J. H. Richter, *Reconfigurable Control of Nonlinear Dynamical Systems: A Fault-Hiding Approach*. Springer, 2011, vol. 408.
- [8] M. M. Seron, J. A. De Dona, and J. H. Richter, "Integrated sensor and actuator fault-tolerant control," *International Journal of Control*, vol. 86, no. 4, pp. 689–708, 2013.

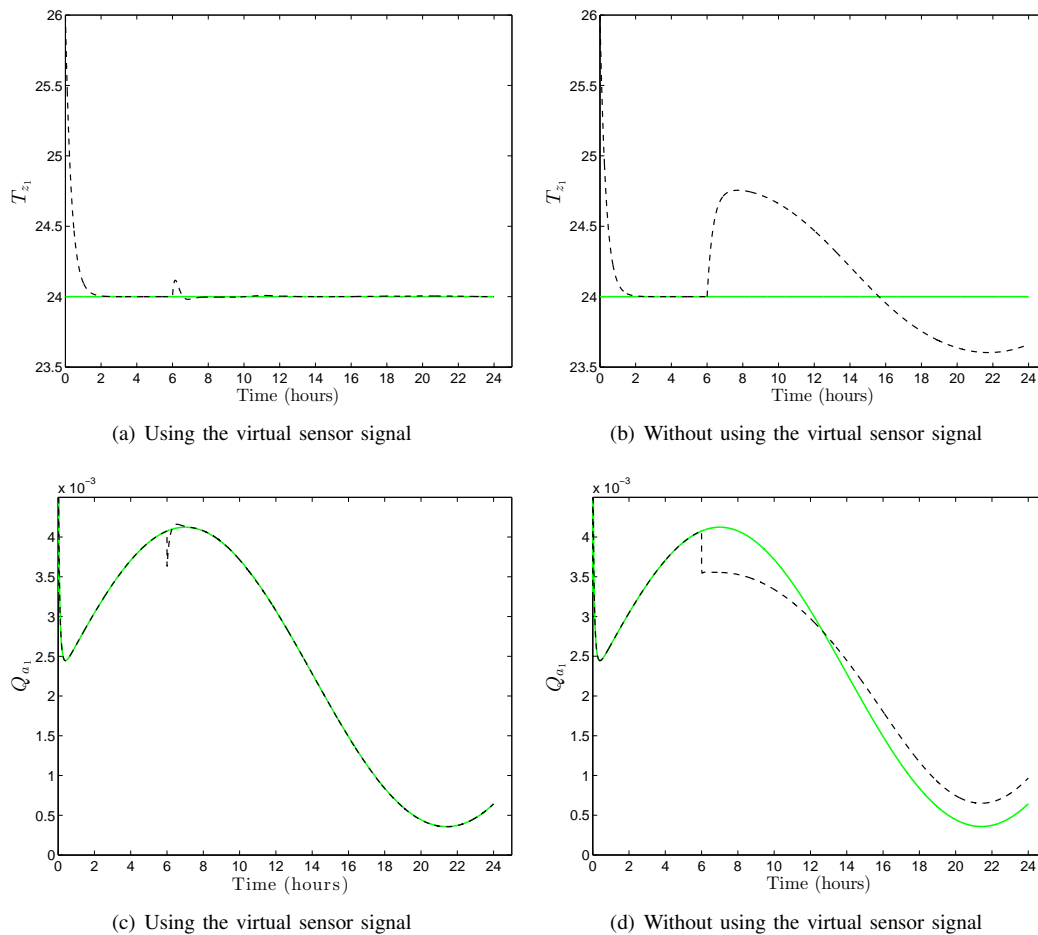


Fig. 4. Effectiveness of the distributed virtual sensor scheme. In (a) and (b), the green line represents the set point of zone temperature, while the black dashed line depicts the zone temperature  $T_{z_1}(t)$  before and after the occurrence of sensor fault in zone one. In (c) and (d), the green line represents the control input  $Q_{a_1}(t)$  under healthy conditions, while the black dashed line corresponds to the control input  $Q_{a_1}(t)$  before and after the occurrence of sensor fault in zone one.

- [9] S. Wang and Y. Chen, "Fault-tolerant control for outdoor ventilation air flow rate in buildings based on neural networks," *Building and Environment*, vol. 37, no. 7, pp. 691–704, 2002.
- [10] X. Hao, G. Zhang, and Y. Chen, "Fault-tolerant control and data recovery in HVAC monitoring system," *Energy and buildings*, vol. 37, no. 2, pp. 175–180, 2005.
- [11] A. Kusiak, M. Li, and H. Zheng, "Virtual models of indoor-air-quality sensors," *Applied Energy*, vol. 87, no. 6, pp. 2087–2094, 2010.
- [12] N. Fernandez, M. R. Brambley, and S. Katipamula, "Self-correcting HVAC controls: Algorithms for sensors and dampers in air-handling units," Pacific Northwest National Laboratory, US Department of Energy, Tech. Rep., 2009.
- [13] V. Reppa, P. Papadopoulos, M. M. Polycarpou, and C. G. Panayiotou, "Distributed detection and isolation of sensor faults in HVAC systems," in *Proceedings of Mediterranean Conference on Control and Automation (MED)*, 2013, pp. 401–406.
- [14] A. Trunov and M. Polycarpou, "Automated fault diagnosis in nonlinear multivariable systems using a learning methodology," *IEEE Transactions on Neural Networks*, vol. 11, no. 1, pp. 91–101, 2000.
- [15] H. A. Talebi, K. Khorasani, and S. Tafazoli, "A recurrent neural-network-based sensor and actuator fault detection and isolation for nonlinear systems with application to the satellite's attitude control subsystem," *IEEE Transactions on Neural Networks*, vol. 20, no. 1, pp. 45–60, 2009.
- [16] B. Thumati and S. Jagannathan, "A model-based fault-detection and prediction scheme for nonlinear multivariable discrete-time systems with asymptotic stability guarantees," *IEEE Transactions on Neural Networks*, vol. 21, no. 3, pp. 404–423, 2010.
- [17] A. Thosar, A. Patra, and S. Bhattacharyya, "Feedback linearization based control of a variable air volume air conditioning system for cooling applications," *ISA Transactions*, vol. 47, pp. 339–349, 2008.
- [18] A. Talukdar and A. Patra, "Dynamic model-based fault tolerant control of variable air volume air conditioning system," *HVAC&R Research*, vol. 16, no. 2, pp. 233–254, 2010.
- [19] J. Farrell and M. Polycarpou, *Adaptive Approximation Based Control*. Wiley-Interscience, 2006.
- [20] V. Reppa, M. M. Polycarpou, and C. G. Panayiotou, "Adaptive approximation for multiple sensor fault detection and isolation of nonlinear uncertain systems," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 137–153, 2014.
- [21] P. Frank, "Handling modelling uncertainty in fault detection and isolation systems," *Journal of Control Engineering and Applied Informatics*, vol. 4, no. 4, pp. 29–46, 2002.
- [22] V. Reppa, M. Polycarpou, and C. G. Panayiotou, "Multiple sensor fault detection and isolation for large-scale interconnected nonlinear systems," in *Proceedings of European Control Conference*, Zurich, Switzerland, 2013, pp. 1952–1957.
- [23] V. Reppa, M. M. Polycarpou, and C. G. Panayiotou, "A distributed detection and isolation scheme for multiple sensor faults in interconnected nonlinear systems," in *52nd Conference on Decision and Control (CDC2013)*, Florence, Italy, 2013, pp. 4991–4996.
- [24] M. Cordier, P. Dague, F. Lévy, J. Montmain, M. Staroswiecki, and L. Travé-Massuyès, "Conflicts versus analytical redundancy relations: a comparative analysis of the model based diagnosis approach from the artificial intelligence and automatic control perspectives," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 34, no. 5, pp. 2163–2177, 2004.