Robust adaptive fault estimation for multizone building VAV terminal units

Mona Subramaniam A, Tushar Jain, and Joseph J. Yamé

Abstract—In this paper, we propose a novel bilinear observer-based robust fault detection, isolation and adaptive fault estimation methodology to precisely estimate a class of actuator faults, namely bias in damper position and lock-in-place faults, in Variable-Air-Volume (VAV) terminal units of Heating Ventilation and Air-Conditioning (HVAC) systems. The proposed adaptive fault estimator is robust in the sense that the fault estimates are not affected by the unmeasured disturbance variable and that the effects of measurement noises on fault estimates are attenuated. The fault estimation algorithm with the integrated building control system improves occupants comfort and reduces the operation, maintenance, and utility cost, thereby reducing the impact on the environment. The effectiveness of the methodology for adaptive estimation of multiple or single VAV damper faults is successfully demonstrated through different simulation scenarios with SIMBAD (SIMulator of Building And Devices), which is being used in industries for testing and validation of building energy management systems.

Note to Practitioners—Faults in VAV dampers of an HVAC system may cause high energy consumption and discomfort to occupants. The complexity of modern HVAC systems, the operational and maintenance costs associated with the equipment necessitate developing an automated techno-commissioning approach by using sophisticated fault diagnosis tools. The fault detection and estimation algorithm presented in this paper does not require linearized models of the building thermal dynamics and information of outside weather conditions. The uniqueness of the developed algorithm is that it is applicable to broader operating conditions and can easily be tailored for robustness with regards to modeling uncertainties and internal heat gains, although the latter can be precisely measured based on the number of occupants. We have shown through exhaustive simulation scenarios that the algorithm detects and estimates the single and multiple occurring stuck and bias faults in VAV dampers, which can be readily integrated with the building energy management system.

Index Terms—Building automation, fault detection, isolation, and estimation, FDD, HVAC-VAV dampers, actuator faults, lock-in-place failure, bias fault, building energy efficiency.

I. INTRODUCTION

Heating Ventilation and Air-Conditioning (HVAC) systems employing Variable Air Volume (VAV) terminal boxes offer an energy-efficient solution to maintain the thermal comfort of occupants inside the building. A VAV box is an end-user terminal unit of the zone, which actuates the volume amount of processed air inside each zone. In order to maintain the desired

thermal comfort inside the zone, a temperature controller in the outer loop generates reference signals to these boxes equipped with electromechanical dampers and internal flow controllers [1]. Several model-based linear [2] and nonlinear [3], and data-driven control [4] techniques are designed for HVAC building systems to ensure the thermal comfort of occupants. However, during the service life of HVAC-VAV units, the sensors and dampers may develop some failures, namely lock-in-place and bias failures. Typically, the detection and identification of the above anomalies involve subunit-by-subunit inspection in a building automation system (BAS) or building energy management system (BEMS), which is challenging to execute frequently. The human operator may not notice this problem when the system is operated in the closed-loop configuration due to a passive control scheme tailored to such faults, where the prime objective is to ensure the tracking of setpoints possibly at the expense of consuming higher energy [5]. As a result, the energy-efficient operation is severely affected under the event of damper lock-in-place and bias faults in VAV boxes. To remove the above manpower and maintenance cost-intensive manual tasks, automated fault detection and diagnosis (FDD) methodologies should be integrated within the building HVAC-VAV systems. This integrated system enables the maintenance crews to identify and predict likely occurring failure modes and to conduct fault impact analysis, determine their criticality on the BAS and be directed in the quick repair process ensuring the system availability. An experimental study on the energy cost impacts for a range of common system faults in various components of the HVAC-VAV systems is conducted in [6].

In the literature, several model-based, data-driven, and grey-box model-based FDD schemes for HVAC-VAV units have been developed [7]. The data-driven or machine learning methods for automated FDD of sensor and actuator faults are developed using neural networks [8], Kalman filters [9], expert knowledge [10], support vector machines (SVM) [11]. Using linear state-observers, model-based FDD methods for detecting and isolating actuator faults have also been developed [12]. Thumati *et al.* [13] presented model-based fault detection and isolation scheme for HVAC-VAV systems for diagnosing degradation faults in the cooling and thermal insulation coils. A linear causal model-based framework for the detection and isolation of sensor and actuator multiple faults in HVAC-VAV systems is presented in [14]. Yoshida *et al.* [15] presented an online FDD in VAV air handling unit to diagnose sensor errors arising out of faulty bias during long term use. The method is based on the dynamic and recursive model building of a VAV control subsystem from the measured data, which

Mona Subramaniam A and Tushar Jain are with Indian Institute of Technology Mandi, School of Computing and Electrical Engineering, Kamand 175075, HP, India e-mail: (d15052@students.iitmandi.ac.in, tushar@iitmandi.ac.in)

Joseph J. Yamé is with Centre de Recherche en Automatique de Nancy, CRAN-CNRS, Université de Lorraine, BP 70239 - 54506 VANDOEUVRE Cedex, France e-mail: (joseph.yame@univ-lorraine.fr)

ensured that robustness against sensor error is self-attained. Darure *et al.* [16] presented a fault estimation algorithm, where a dedicated bank of unknown input residual generators was employed for fault detection and isolation (FDI). The nonlinear observer-based schemes for HVAC sensor FDI is presented in [17]. In our previous work [18], we have also developed a bilinear observer-based complete fault detection, isolation, and accurate estimation of the magnitude of actuator faults. In the above data-driven and model-based FDD schemes, the robustness of the fault diagnosis methods is not studied. The robustness problem of an FDD method deals with the accuracy of detection, isolation, and estimation of the magnitude of faults under several types of uncertainties, such as modelling errors, sensor measurement noise, transients due to switching of operating mode and unknown disturbances. Wang *et al.* [19] presented a robust FDD strategy for the isolation of faults in pressure-independent VAV dampers using a hybrid approach based on cumulative sum control charts and a rule-based fault classifier. The hybrid approach is robust to false alarms caused by impacts of normal transient changes within the system. Bonvini *et al.* [20] presented a robust FDD algorithm based on Unscented Kalman Filter techniques for detecting common chiller plant faults in an HVAC system, which is robust to sensor errors and data paucity. Mulumba *et al.* [21] proposed a robust model-based fault diagnosis method based on SVM techniques, which is robust to model parameter uncertainty. The effectiveness of another hybrid approach based on the SVM technique for fault detection and isolation is demonstrated through numerical simulation in [22]. Zhao *et al.* [23] presented a comprehensive survey on artificial intelligence-based FDD methods concluding that the effectiveness of data-driven and knowledge-based schemes rely heavily on training data and expert knowledge respectively, without which robustness cannot be achieved.

With regards to the vast development of fault diagnosis approaches for HVAC-VAV-based buildings, many of them are built upon linear and data-driven statistical models of the building thermal dynamics. Since a fault diagnosis module is an integral part of a control scheme, addressing the aftereffects of faults, statistical or purely data-based techniques may not be useful in implementing an active fault-tolerant control (AFTC) scheme [24]. With the dynamically changing zone occupancy profile and outside weather conditions, the linearized models capturing the zone thermal dynamics need to be updated at regular time intervals in linear model-based approaches. To address this issue, an observer-based fault diagnosis approach is developed in our previous work in [18] by dealing directly with the true bilinear thermal dynamics. The fault diagnosis problem for control affine nonlinear process systems has also been addressed in [25] and in the presence of uncertainty in [26]. However, in these works, only the fault detection and isolation modules are developed using high-gain observers for at most two simultaneously occurring faults under the absence of measurement noise. The robustness issues with regards to unmeasurable disturbances and measurement noise remains an open problem. The main contributions of this paper are:

- a detailed design of a novel fault diagnosis methodology

integrated with robust adaptive estimation of faults, which is inherited from the true nonlinear (bilinear) dynamics of controlled thermal zones by VAV terminal units is derived, thereby making the methodology applicable to a broad operating range of the system;

- the robustness of the algorithm with regards to unmeasured disturbances and noisy sensor measurements is rigorously established such that the fault estimates are insensitive to unmeasured disturbances, and the effect of measurement noise on these estimates is attenuated;
- a wide class of actuator faults, namely lock-in-place faults and bias faults in VAV dampers, is considered in this paper;
- an extensive simulation study of a one-storey building with three-zones is presented within the Matlab-based SIMBAD (SIMulator of Building And Devices) software environment to demonstrate the effectiveness of the derived algorithms, where two scenarios consisting of four different single and multiple VAV damper fault cases are considered.

The paper is organized as follows: The system description, fault modelling, and the problem statement are given in Section II. In Section III, the robust adaptive fault estimation algorithm based on a nonlinear observer is derived. The two different simulation scenarios composed of four different cases of single and multiple faults occurring in HVAC-VAV dampers of a one-storey three-zone building are presented in Section IV. Finally, in Section V, some concluding remarks are provided.

II. SYSTEM DESCRIPTION AND PROBLEM FORMULATION

In this section, firstly, we briefly present the bilinear model of the thermal dynamics in a multizone building. The detailed modelling equations can be referred to [18, Section 2]. Subsequently, two types of VAV damper faults are modelled, followed by the problem statement addressed in this paper.

A. Thermal dynamics of a multizone building

The air conditioning technical plan of a typical HVAC-VAV system in a multizone building with its traditional control loops is sketched in Fig. 1. Each zone of the building is actuated by a VAV unit receiving the processed air from the central air-handling unit (AHU) at a constant supply air temperature, T_s . The AHU being the centralized subunit, includes a mixer, cooling/heating coils, and a variable-speed supply fan. The air is processed within the AHU subunit through the mixer damper, mixing the outside air with the return air from all the zones, and cooling/heating coils, controlling the supply airflow temperature, which is separately controlled by conventional/non-conventional energy sources. Each VAV box is equipped with a controlled damper that regulates the volume of hot/cold primary air depending on the outside weather conditions to be delivered inside the zones according to thermal requirements. Under the standard modelling assumptions given

Fig. 1. Technical layout of the HVAC-VAV system in a multizone building and its traditional cascaded control strategy

TABLE I
BUILDING THERMAL VARIABLES AND PARAMETERS

Variables	Description
T_{zi}	Temperature of the i -th Zone
T_z	Zone temperature vector
$A, B, C, \mathcal{D}, E$	Building thermal parameters
$F(t)$	VAV damper fault distribution matrix
$\dot{m}_i = u_i$	Airflow rate setpoints to the i -th zone
$Q_{int,i}$	Internal heat gain of the i -th zone
Q_{int}	Internal gain vector
T_{oa}	Outside air ambient temperature
$T_{extw,i}$	Wall temperature of the i -th zone external wall.
$T_{intw,ij}$	Wall temperature of the internal wall between the i -th and j -th zone
T_{extw}	External wall temperature vector
T_{intw}	Internal wall temperature vector
T_w	Wall temperature vector
T_s	Supply air temperature
η	Measurement noise signal vector
f_i	Fault in the i -th zone
f_a	Fault vector containing faults of all zones
u_f	Actual airflow rate to the zones
L, R	Fault estimator design matrices
κ_{0i}	Constant coefficient parameters of the i -th zone building thermal dynamics
$\vartheta(t)$	impulse response of the causal filter
e_i	i -th column vector of the standard basis of $\mathbb{R}^n$

with

$$\begin{aligned}
 T_z &= [T_{z1} \ \cdots \ T_{zn}]^T, \quad T_w = [T_{extw}^T \mid T_{intw}^T]^T \\
 T_{extw} &= [T_{extw,1} \ \cdots \ T_{extw,p_1}]^T \\
 T_{intw} &= [T_{intw,12} \ \cdots \ T_{intw,ij}]^T \\
 Q_{int} &= [Q_{int,1} \ \cdots \ Q_{int,n}]^T
 \end{aligned} \tag{3}$$

for $i, j = 1, 2, \dots, n, ij \neq ji$, model parameters $\mathcal{D}, A, B, C, E$ are matrices of compatible dimensions, and p_1 being the dimension of external walls. These matrices are composed of thermodynamical parameters, which can be modelled as electrical quantities in terms of capacitances and thermal resistances of the zones and that of the external and internal walls [18]. The description of the variables introduced in (2)-(3) is provided in Table I. Let $p = p_1 + p_2$ be the total number of walls, with $T_{extw} \in \mathbb{R}^{p_1}$ and $T_{intw} \in \mathbb{R}^{p_2}$, then $x \in \mathbb{R}^{n+p}$, $u \in \mathbb{R}^n$, $y \in \mathbb{R}^n$, and $d \in \mathbb{R}^{n+1}$. The output matrix $C = [I_n \ 0_{n \times p}]$, where I_n is the n -dimensional identity matrix, indicates that the zone temperatures are measured directly but the wall temperatures. The bilinear term can be expressed as

$$\mathcal{D}(u \otimes x) = \sum_{i=1}^n D_i \dot{m}_i(t) x(t) = D_u(t) x \tag{4}$$

in [18], the following bilinear state-space model describes the thermal dynamics of multizone buildings

$$\begin{bmatrix} \dot{x} \\ y \end{bmatrix} = \begin{bmatrix} \mathcal{D}(u \otimes x) \\ 0_{n \times 1} \end{bmatrix} + \begin{bmatrix} A & B & E \\ C & 0_{n \times n} & 0_{n \times (n+1)} \end{bmatrix} \times \begin{bmatrix} x^T & u^T & d^T \end{bmatrix}^T \tag{1}$$

where ' $\otimes$ ' denotes the Kronecker product, $0_{\bullet}$ denotes the zero matrix of appropriate dimension, x, u, y , and d denote the state, input, output and the disturbance variable given as

$$\begin{aligned}
 x(t) &= [T_z^T \mid T_w^T]^T, \quad y(t) = T_z, \\
 u(t) &= [\dot{m}_1 \ \cdots \ \dot{m}_n]^T, \quad d(t) = [Q_{int}^T \mid T_{oa}]^T,
 \end{aligned} \tag{2}$$

with $\mathcal{D} = [D_1 \ \cdots \ D_n]$, where D_i 's are square matrices of order $(n + p)$ and $D_u(t)$ is a time-varying matrix of appropriate dimension. Typically, sensors are deployed to measure the zone temperatures, which inherently induce noise in the measured output signals. The sensor measurement noise can be modelled by a uniformly distributed additive noise signal denoted by $\eta(t)$. Then, the governing overall state-space model is represented by

$$(\Sigma) : \begin{cases} \begin{bmatrix} \dot{x} \\ z \end{bmatrix} = \begin{bmatrix} \mathcal{D}(u \otimes x) \\ 0_{\bullet} \end{bmatrix} + \begin{bmatrix} A & B & E \\ C & 0_{\bullet} & 0_{\bullet} \end{bmatrix} \times \begin{bmatrix} x^T & u^T & d^T & \eta^T \end{bmatrix}^T \end{cases} \tag{5}$$

where signal $z(t)$ is the measured output.

B. Fault modelling in VAV terminal units

To ensure the energy-efficiency of the building HVAC system while satisfying the thermal comfort requirement within the zones, the fault-free operation of VAV dampers plays a significant role because based on the thermal load, these dampers adjust the volume flow rate of processed air being dispatched into the zone. The outer-loop temperature controller ensures that the temperature of each zone is regulated to its setpoint by providing the variable airflow rate setpoint to the inner-loop flow controller. Whenever the VAV damper of the i -th zone becomes faulty, the energy-efficient operation of the HVAC system and the thermal comfort of the occupants of the building are drastically hampered. Two types of frequently occurring VAV damper faults, namely a constant bias in the damper position and damper lock-in-place failure, are considered in this work. A lock-in-place failure occurs when the VAV damper is stuck anywhere between the operating range of the VAV damper given by $[0, 1]$, where 0 represents a damper fully closed, and 1 when it is fully open. A bias fault refers to the case when the commanded position by the controller is offset by a bias in the actual damper position. Let $f_a(t)$ denote the actuator fault signal.

- **Fault type - 1:** lock-in-place of VAV dampers: Under the event of “fault type - 1”, the input signal in (5) is replaced by the faulty signal, $u_f(t)$, modelled as

$$u_f(t) = [\mathbf{1}_n^T \otimes I_n] [u(t)^T \mid f_a^T(t)]^T \quad (6)$$

such that $u_f(t)$ is a constant signal, where $\mathbf{1}_n$ is the n -dimensional vector of entries equal to 1.

- **Fault type - 2:** constant bias in VAV damper position: Under the event of “fault type - 2”, the fault can also be modelled as given in (6), but $u_f(t)$ may not be a constant signal with $f_a(t)$ denoting a constant bias [24].

Substituting any of the above fault types in (Σ) yields the following state-space model of the faulty system as

$$(\Sigma_f) : \begin{cases} \begin{bmatrix} \dot{x} \\ z \end{bmatrix} = \begin{bmatrix} A & B & E & 0_\bullet & F(t) \\ C & 0_\bullet & 0_\bullet & I_\bullet & 0_\bullet \end{bmatrix} \\ \times \begin{bmatrix} x^T & u^T & d^T & \eta^T & f_a^T \end{bmatrix}^T \\ + \begin{bmatrix} \mathcal{D}(u \otimes x) \\ 0_\bullet \end{bmatrix} \end{cases} \quad (7)$$

where $f_a = [f_1 \cdots f_n]^T$ denotes the fault vector. Let denote $\kappa = \text{diag}(\kappa_{01}, \dots, \kappa_{0n})$, see Table I, with $\text{diag}(\bullet)$ denoting the diagonal matrix of $\bullet$ element, and e_i be the i -th column vector of the standard basis of $\mathbb{R}^n$, then $F(t) = [F_1(t) \ 0_{p \times n}]$, with $F_1(t) = \kappa (\sum_{i=1}^n e_i e_i^T (\mathbf{1}_n T_s - y(t)) e_i e_i^T)$ a time-varying fault distribution matrix.

C. Problem formulation

The problem we aim to address in this paper is to design a bilinear observer for the robust adaptive estimation of the magnitude of VAV damper faults under the effect of partially

decouplable disturbances such that the energy-efficient operation of the overall closed-loop system can be maintained. To this end, the disturbance signal, $d(t)$, in (5) is partitioned into measured and unmeasured disturbance signals denoted by $d_m(t)$ and $d_{um}(t)$ respectively, i.e., $d = [d_m^T \mid d_{um}^T]^T$. The associated matrix E can also be partitioned as $E = [E_m \mid E_{um}]$. Taking into account the above partition, the objective is to ensure the robustness of the fault estimator with regards to the effect of the unmeasured disturbance $d_{um}(t)$ and that the effects of $\eta(t)$ are attenuated as much as possible. In the literature, several control schemes are reported assuming the availability of either both the disturbances [27] or one of them while estimating the other [28]. In [29], moving average models are proposed to estimate both the disturbances in real-time for developing the model-predictive control (MPC) scheme. In [30], the measurement of outside weather conditions is assumed to be known, while the information of occupancy and subsequently, the internal heat gain is estimated online in the MPC scheme. A practical implementation of the MPC scheme assuming the availability of both disturbances is developed in [31], [32]. In [33], a simultaneous identification method is proposed to identify both linear thermal model (of resistance-capacitance network type) and the internal load from measured input-output data assuming that the latter is constant and therefore modeled as a process with a zero derivative. In our previous work [18], the information of both disturbances was assumed to be known in the solution of the fault estimation problem for HVAC building systems. For its practical implementation, this may additionally require high computational modules to precisely estimate these disturbances in real-time [34]. To address this issue, the solution to the fault estimation problem should be robust with regards to these disturbances. It is well known in the literature that such robust solutions are constrained by some matrix existence conditions, which may not be satisfied in some scenarios [35]. This work takes further steps towards a realistic approach by assuming that the disturbance is only known partially. Particularly, the following partition of the disturbance is used in the rest of the text $d_m = Q_{int}$, and $d_{um} = T_{oa}$. Note that the former disturbance can be computed online using either stochastic models [29] or a technique based on passive infrared sensors [32]. The latter disturbance is a weather variable which significantly impacts the building dynamics as an unknown large disturbance input. Clearly, the inherent uncertainty in the weather and hence in the outdoor air temperature is a major challenge that requires a proper treatment in the design of FDD algorithms for building HVAC systems. The main non-trivial issues to be solved for the fault estimation problem are, therefore:

- How to detect and isolate the faults?
- How to estimate the fault magnitude?
- How to remove the effect of unmeasured disturbance on fault estimation?
- How to attenuate the effect of sensor measurement noise on fault estimation?

III. ROBUST ADAPTIVE FAULT ESTIMATOR

The following definition and lemma are required to develop the fault estimation algorithm.

Definition 1: [36], [37] The dynamical system

$$\begin{bmatrix} \dot{x} \\ y \end{bmatrix} = \begin{bmatrix} A & B & E \\ C & 0_{\bullet} & 0_{\bullet} \end{bmatrix} \begin{bmatrix} x^T & u^T & d^T \end{bmatrix}^T + \begin{bmatrix} \phi(u, y) \\ 0_{\bullet} \end{bmatrix}$$

with $\phi(u, y)$ representing a matrix-valued nonlinear function of input and output variables is defined as an output injected nonlinear system.

Proposition 1: The dynamical system (Σ) is an output injected nonlinear system.

Proof: The bilinear term in (Σ) reads as $\mathcal{D}(u \otimes$

$$x) = \phi(u, x) = \begin{bmatrix} D_{11}(t) & D_{12}(t) \\ D_{21}(t) & D_{22}(t) \end{bmatrix} \begin{bmatrix} T_z \\ T_w \end{bmatrix} \text{ where } D_{11}, D_{12}, D_{21}, \text{ and } D_{22} \text{ are time-varying block matrices of } D_u(t), \text{ see (4). Thanks to the model structure of the multizone building thermal dynamics, matrices } D_{12}(t), D_{21}(t), D_{22}(t) \text{ are all zero matrices of an appropriate dimension. This results in}$$

$$\begin{aligned} \mathcal{D}(u \otimes x) &= \phi(u, x) = \begin{bmatrix} D_{11}(t) & 0_{n \times p} \\ 0_{p \times n} & 0_{p \times p} \end{bmatrix} \begin{bmatrix} T_z \\ T_w \end{bmatrix} \\ &= \begin{bmatrix} D_{11}(t)T_z \\ 0_{p \times 1} \end{bmatrix} = \begin{bmatrix} D_{11}(t)y \\ 0_{p \times 1} \end{bmatrix} = \phi(u, y). \end{aligned} \quad (8)$$

Moreover, $D_{11}(t) = -\kappa (\sum_{i=1}^n e_i e_i^T u(t) e_i e_i^T)$. Clearly, (Σ) is an output injected nonlinear (bilinear) system. ■

A. Development of the VAV damper fault estimation algorithm

The bilinear observer-based robust adaptive fault estimator (RAFE) is proposed as

$$\begin{aligned} \begin{bmatrix} \dot{\hat{x}} \\ \hat{y} \\ \hat{f}_a \end{bmatrix} &= \begin{bmatrix} A - LC & B & E_m & L & F(t) \\ C & 0_{\bullet} & 0_{\bullet} & 0_{\bullet} & 0_{\bullet} \\ -(\vartheta(t) * GC) & 0_{\bullet} & 0_{\bullet} & (\vartheta(t) * G) & 0_{\bullet} \end{bmatrix} \\ &\quad \times \begin{bmatrix} \hat{x}^T & u^T & d_{um}^T & z^T & \hat{f}_a^T \end{bmatrix}^T \\ &\quad + \begin{bmatrix} \phi(u, z)^T & 0_{\bullet} & 0_{\bullet} \end{bmatrix}^T \end{aligned} \quad (9)$$

where $*$ represents the convolution operator, $\hat{x}(t)$ denotes the state estimate, $\hat{y}(t)$ the output estimate. Here, it is worth noticing that the unknown fault $f_a(t)$ is viewed as a time-varying parameter whose estimate $\hat{f}_a(t)$ is obtained by the adjusting law given by $\hat{f}_a(t)$ in (9), where the to-be-designed parameters are constant matrices L and G , and $\vartheta(t)$ the impulse response of a causal filter. Let the state estimation error be defined as $\varepsilon(t) = x(t) - \hat{x}(t)$, then from (Σ_f) and (9) the error dynamics reads

$$\begin{aligned} \dot{\varepsilon} &= \begin{bmatrix} (A - LC) & E_{um} & -L & F(t) \end{bmatrix} \\ &\quad \times \begin{bmatrix} \varepsilon^T & d_{um}^T & \eta^T & (f_a - \hat{f}_a)^T \end{bmatrix}^T \\ &\quad + (\phi(u, y) - \phi(u, z)) \end{aligned} \quad (10)$$

Proposition 2: Define $\tilde{\eta}(t)$ as $\tilde{\eta} = [\eta_u^T \ \eta^T]^T$ with $\eta_u(t) = (\sum_{i=1}^n e_i e_i^T u(t) e_i e_i^T) \eta(t)$. Then, the nonlinear error dynamics (10) is equivalent to the linear dynamics

$$\begin{aligned} \dot{\varepsilon} &= \begin{bmatrix} \tilde{A} & \tilde{B} & E_{um} & F(t) \end{bmatrix} \\ &\quad \times \begin{bmatrix} \varepsilon^T & \tilde{\eta}^T & d_{um}^T & (f_a - \hat{f}_a)^T \end{bmatrix}^T \end{aligned} \quad (11)$$

where $\tilde{A} = (A - LC)$, and $\tilde{B} = \begin{bmatrix} \frac{1}{T_s} B & -L \end{bmatrix}$.

Proof: Using Proposition 1, the matrix-valued nonlinear function $\phi(u, z)$ in (10) can be decomposed as the sum of $\phi(u, y)$ and a matrix consisting of the product of $D_{11}(t)$ and η , i.e.,

$$\begin{aligned} \phi(u, z) &= \phi(u, y + \eta) = \begin{bmatrix} D_{11}(t)(y + \eta) \\ 0_{p \times 1} \end{bmatrix} \\ &= \phi(u, y) + \begin{bmatrix} D_{11}(t)\eta \\ 0_{p \times 1} \end{bmatrix} \end{aligned} \quad (12)$$

where matrix $D_{11}(t)\eta$ can be written as $D_{11}(t)\eta = -\kappa (\sum_{i=1}^n e_i e_i^T u(t) e_i e_i^T) \times \eta(t) = -\kappa \eta_u(t)$. Putting (12) into (10) removes the nonlinear term from the error dynamics. Thanks to the model structure of a multizone building, the matrix B is given as $B = \begin{bmatrix} B_1^T & 0_{\bullet}^T \end{bmatrix}^T$ where $B_1 \in \mathbb{R}^{n \times n}$ with $B_1 = \kappa T_s$. As a consequence, the second term in (12) can be expressed as $\begin{bmatrix} D_{11}(t)\eta \\ 0_{p \times 1} \end{bmatrix} = -\frac{1}{T_s} B \eta_u$. which on substituting in (10) yields the linear system (11). ■

For fault estimation, we use the error dynamics resulting from a similarity transformation $\varepsilon = \Gamma \zeta$ applied on (11), where matrix Γ is computed such that $\tilde{A}$ is a diagonal matrix [38]. Consequently, (11) is transformed to

$$\begin{aligned} \begin{bmatrix} \dot{\zeta} \\ \zeta_y \end{bmatrix} &= \begin{bmatrix} \tilde{A} & \tilde{B} & \tilde{E}_{um} & \tilde{F}(t) & 0_{\bullet} \\ \tilde{C} & 0_{\bullet} & 0_{\bullet} & 0_{\bullet} & I_{\bullet} \end{bmatrix} \\ &\quad \times \begin{bmatrix} \zeta^T & \tilde{\eta}^T & d_{um}^T & (f_a - \hat{f}_a)^T & \eta^T \end{bmatrix}^T \end{aligned} \quad (13)$$

where $\zeta_y = z - \hat{y}$, $\tilde{A} = \Gamma^{-1} \tilde{A} \Gamma$, $\tilde{B} = \Gamma^{-1} \tilde{B}$, $\tilde{E}_{um} = \Gamma^{-1} E_{um}$, $\tilde{F}(t) = \Gamma^{-1} F(t)$ and $\tilde{C} = C \Gamma$. Premultiplying the derivative of the second block vector in (13) by some nonzero matrix G , and substituting the first block vector in the resulting vector yields

$$\begin{aligned} G \dot{\zeta}_y &= \begin{bmatrix} G \tilde{C} \tilde{A} & G \tilde{C} \tilde{B} & G \tilde{C} \tilde{E}_{um} & G \tilde{C} \tilde{F}(t) & G \end{bmatrix} \\ &\quad \times \begin{bmatrix} \zeta^T & \tilde{\eta}^T & d_{um}^T & (f_a - \hat{f}_a)^T & \eta^T \end{bmatrix}^T \end{aligned} \quad (14)$$

From (14), vector $G \dot{\zeta}_y$ is not affected by the unmeasured disturbance if the signal $\tilde{C} \tilde{E}_{um} d_{um} \in \ker(G)$, where $\ker(\bullet)$ denotes the kernel of a matrix. The following proposition is one of the key results of the paper and provides a sufficient condition for the existence of a matrix G .

Proposition 3: If $\tilde{C} \tilde{E}_{um} = [h_1 \ \dots \ h_n]^T$ with $h_i \neq 0, i = 1, 2, \dots, n$, then there exists a gain matrix G such that

- i) the vector $\tilde{C} \tilde{E}_{um} d_{um} \in \ker(G)$, and
- ii) $\tilde{C} \tilde{F}(t)(f_a - \hat{f}_a) \notin \ker(G)$.

Proof: The proof of the proposition is constructive. Indeed, construct matrix $G \in \mathbb{R}^{n \times n}$ as

$$G = \begin{bmatrix} 1 & -\frac{h_1}{h_2} & 0 & \cdots & 0 & 0 \\ 0 & 1 & -\frac{h_2}{h_3} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -\frac{h_{n-1}}{h_n} \\ -\frac{h_n}{h_1} & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

- i) then it is straightforward to check that $\bar{C}\bar{E}_{um}d_{um} \in \ker(G)$.
 ii) From (7), since $F(t) = \begin{bmatrix} \kappa(\sum_{i=1}^n e_i e_i^T (\mathbf{1}_n T_s - T_z) e_i e_i^T) & \mathbf{0}_\bullet \end{bmatrix}$, the matrix $GCF(t)$ is expressed as $G\bar{C}\bar{F}(t) = G \times \text{diag}(F_{11}(t), \dots, F_{nn}(t))$ where $F_{ii}(t)$ denotes the i -th diagonal element of $F(t)$. Writing explicitly the above matrix, we get

$$G\bar{C}\bar{F}(t) = \begin{bmatrix} F_{11} & -\frac{h_1}{h_2}F_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{h_n}{h_1}F_{11} & 0 & \cdots & F_{nn} \end{bmatrix}. \quad (15)$$

Now the vector $\bar{C}\bar{F}(t)(f_a - \hat{f}_a) \in \ker(G)$ if and only if $T_s \mathbf{1}_n - T_z = \mathbf{0}_n$. Thanks to the building HVAC dynamics, the last condition is never satisfied.

Remark 1: The rank of the matrix G and consequently, that of $G\bar{C}\bar{F}(t)$ is $n-1$ for all time t , which implies that at least one of the eigenvalues of these matrices is zero. From the structure of the matrix G , there is only one eigenvalue at zero. Since the rank deficient matrix G is used in the adjusting law $\dot{\hat{f}}_a(t)$, only $n-1$ faults can be estimated from (9). Nevertheless, multiple VAV damper faults can also be estimated using a multiple fault estimator scheme based on (9), which shall be presented in the next section.

Remark 2: The design matrix G in Proposition 3 is well defined whenever $h_i \neq 0$, for all $i = 1, 2, \dots, n$. It is worth noting that $\bar{C}\bar{E}_{um} = CE_{um}$. Since $C = \begin{bmatrix} I_n & \mathbf{0}_{n \times p} \end{bmatrix}$, and the outside ambient temperature always affects the zone dynamics, every element of the associated distribution vector is nonzero.

We are now in a position to state the main result of the paper in the next proposition.

Proposition 4: Choose the causal filter $\vartheta(t)$ as

$$\vartheta(t) = \begin{bmatrix} \beta_1 & \beta_2 \end{bmatrix} \begin{bmatrix} \delta(t) & \mathbb{1}(t) \end{bmatrix}^T \quad (16)$$

where $\beta_1, \beta_2 \in \mathbb{R}^{n \times n}$ are the filter design matrices with $\delta(t)$ and $\mathbb{1}(t)$ denoting the Dirac delta function and the Heaviside function respectively. If

- the matrix β_1 is selected such that $\bar{C}\bar{F}(t)\tilde{f}_a \notin \ker(\beta_1 G)$, where $\tilde{f}_a = f_a - \hat{f}_a$ denotes the fault estimation error, and

- the matrices β_2 and L are designed such that the matrix $\begin{bmatrix} A - LC & B \\ -\beta_2 GC & \mathbf{0}_\bullet \end{bmatrix}$ is Hurwitz,

then the fault estimation error is independent of the unmeasured disturbance signal d_{um} and the effect of measurement noise η on $\hat{f}_a$ is attenuated.

Proof: Substituting $z(t)$ from (7) in the adjusting law $\dot{\hat{f}}_a$ in the last block-row in (9) and using transformation (13) yields

$$\begin{aligned} \dot{\hat{f}}_a(t) = & (\beta_1 G\bar{C}\bar{A} + \beta_2 G\bar{C})\zeta + \beta_1 G\bar{C}\bar{B}\tilde{\eta} + \dot{\vartheta}(t) * G\eta \\ & + \beta_1 G\bar{C}\bar{E}_{um}d_{um} + \beta_1 G\bar{C}\bar{F}(t)\tilde{f}_a \end{aligned} \quad (17)$$

From Proposition 3, the fourth term in (17) is zero. The solution of the linear system (13) is computed as

$$\begin{aligned} \zeta(t) = & e^{\bar{A}(t-t_0)}\zeta(t_0) + \int_{t_0}^t e^{\bar{A}(t-\tau)} (\bar{B}\tilde{\eta}(\tau) + \bar{E}_{um}d_{um}(\tau) \\ & + \bar{F}(\tau)(f_a(\tau) - \hat{f}_a(\tau))) d\tau. \end{aligned} \quad (18)$$

Putting (18) into (17) gives

$$\begin{aligned} \dot{\hat{f}}_a = & (\beta_1 G\bar{C}\bar{A} + \beta_2 G\bar{C}) \int_{t_0}^t e^{\bar{A}(t-\tau)} (\bar{E}_{um}d_{um}(\tau) \\ & + \bar{B}\tilde{\eta}(\tau)) d\tau + \beta_1 G\bar{C}\bar{B}\tilde{\eta} + \dot{\vartheta}(t) * G\eta + \text{o.t.}, \end{aligned} \quad (19)$$

where in other terms (o.t.) there is no signature of $d_{um}(t)$. Since the matrix $\bar{A}$ is diagonal, the associated matrix exponential is also diagonal. It is now clear from (19) that if $\bar{C}\bar{E}_{um}d_{um} \in \ker(\beta_1 G + \beta_2 G)$, then there is no signature of $d_{um}(t)$ in $\dot{\hat{f}}_a(t)$ and if the matrix β_1 is selected such that $\bar{C}\bar{F}(t)\tilde{f}_a \notin \ker(\beta_1 G)$, then the fault signature is present in $\dot{\hat{f}}_a(t)$. From Proposition 3, $\bar{C}\bar{E}_{um}d_{um} \in \ker(G)$ implies $\bar{C}\bar{E}_{um}d_{um} \in \ker(\beta_1 G + \beta_2 G)$.

Substituting $\bar{B}$ and $\tilde{\eta}$ from Proposition 2 in equation (17), the derivative of the fault estimation error $\tilde{f}_a$ reads

$$\begin{aligned} \dot{\tilde{f}}_a = & -(\beta_1 G\bar{C}\bar{A} + \beta_2 G\bar{C})\zeta - \frac{1}{T_s} \beta_1 G\bar{C}\Gamma^{-1} B\eta_u + \dot{f}_a \\ & + \beta_1 G\bar{C}\Gamma^{-1} L\eta - \dot{\vartheta}(t) * G\eta - \beta_1 G\bar{C}\bar{F}(t)\tilde{f}_a, \end{aligned} \quad (20)$$

and the state equation of (13) reduces to

$$\dot{\zeta} = \bar{A}\zeta + \frac{1}{T_s} \Gamma^{-1} B\eta_u - \Gamma^{-1} L\eta + \bar{E}_{um}d_{um} + \bar{F}(t)\tilde{f}_a. \quad (21)$$

If there is a damper lock-in-place failure in VAV boxes, then $u_f(t)$ is a constant signal. In addition, under closed-loop control, the airflow setpoint generated by the controller, $u(t)$, saturates due to hard constraints on the VAV dampers. From (6), if $u(t)$ and $u_f(t)$ are constant signals under the fault type-2, then $f_a(t)$ is also a constant signal and consequently, $\dot{f}_a(t) = 0$. For fault type-2, $\tilde{f}_a(t)$ is also equal to zero. Con-

catenating vectors ζ and $\dot{\hat{f}}_a$, the derivative of the concatenated vector reads (see equations (20) and (21))

$$\begin{aligned} \begin{bmatrix} \dot{\zeta} \\ \dot{\hat{f}}_a \end{bmatrix} &= \frac{1}{T_s} \begin{bmatrix} \Gamma^{-1}B \\ -\beta_1 G \bar{C} \Gamma^{-1} B \end{bmatrix} \eta_u + \begin{bmatrix} \bar{E}_{um} \\ 0_{\bullet} \end{bmatrix} d_{um} \\ &+ \begin{bmatrix} \bar{A} & \bar{F}(t) \\ -(\beta_1 G \bar{C} \bar{A} + \beta_2 G \bar{C}) & -\beta_1 G \bar{C} \bar{F}(t) \end{bmatrix} \begin{bmatrix} \zeta \\ \hat{f}_a \end{bmatrix} \\ &+ \begin{bmatrix} -\Gamma^{-1}L \\ \beta_1 G \bar{C} \Gamma^{-1} L - \dot{\eta}(t) * G \end{bmatrix} \eta. \quad (22) \end{aligned}$$

Set $\check{A}(t) = \begin{bmatrix} \bar{A} & \bar{F}(t) \\ -(\beta_1 G \bar{C} \bar{A} + \beta_2 G \bar{C}) & -\beta_1 G \bar{C} \bar{F}(t) \end{bmatrix}$, and define a unitriangular matrix $\Lambda = \begin{bmatrix} I_{n+p} & 0_{(n+p) \times n} \\ \beta_1 G \bar{C} & I_n \end{bmatrix}$.

Premultiplying $\check{A}(t)$ with Λ yields

$$\Lambda \check{A}(t) = \begin{bmatrix} \bar{A} & \bar{F}(t) \\ -\beta_2 G \bar{C} & 0_{\bullet} \end{bmatrix}. \quad (23)$$

Thanks to the model structure of the building thermal dynamics, the time-varying fault distribution matrix $F(t)$ can be expressed as $F(t) = B + H_z(t)$ where $H_z(t) = [H_{z1}^T(t) \ 0_{\bullet}^T]^T$ with $H_{z1}(t) = -\kappa (\sum_{i=1}^n e_i e_i^T y(t) e_i e_i^T)$, then $\bar{F}(t) = \Gamma^{-1} F(t) = \Gamma^{-1} (B + H_z(t))$. Substituting $\bar{F}(t)$ and $\bar{A} = \Gamma^{-1} (A - LC) \Gamma$ in (23), the matrix $\Lambda \check{A}(t)$ can be decomposed as the sum of a constant matrix $\mathbf{A}$ and a time-varying matrix $\mathcal{A}(t)$, i.e., $\Lambda \check{A}(t) = \mathbf{A} + \mathcal{A}(t)$, where

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} \Gamma^{-1} & 0_{\bullet} \\ 0_{\bullet} & I_{\bullet} \end{bmatrix} \begin{bmatrix} A - LC & B \\ -\beta_2 G \bar{C} & 0_{\bullet} \end{bmatrix} \begin{bmatrix} \Gamma & 0_{\bullet} \\ 0_{\bullet} & I_{\bullet} \end{bmatrix}, \\ \mathcal{A}(t) &= \begin{bmatrix} 0_{\bullet} & \Gamma^{-1} H_z(t) \\ 0_{\bullet} & 0_{\bullet} \end{bmatrix}. \end{aligned}$$

Now the fault estimator (9) matrices L and β_2 are computed such that the matrix $\mathbf{A}$ or equivalently $\begin{bmatrix} A - LC & B \\ -\beta_2 G \bar{C} & 0_{\bullet} \end{bmatrix}$ is Hurwitz. Moreover the time-varying matrix $\mathcal{A}(t)$ is always bounded since it is a function of the zone temperatures, which remains bounded due to operating limits on the processed airflow rate. Therefore, the zero input dynamics of the system defined by (22) is asymptotically stable. In addition, the signals (u, η) to (22) are bounded, the fault estimation error is also bounded. ■

Remark 3: The main focus of this paper is on the issue of FDD robustness against uncertainties arising from noise and unmeasured disturbances in building equipped with VAV terminal units subject to a wide class of faults. It is clear that, as the FDD algorithm is based on the building dynamical model, there will be always a mismatch between the actual building and its mathematical model even in the fault-free case. Such discrepancies can be sources of false and missed alarms which can compromise the FDD system performance. To overcome such difficulties, the FDD algorithm must be enhanced with robustness against uncertainties arising from the building modeling. Note that for lock-in-place failures of

the VAVs, the diagnosis performed by the algorithm is virtually unaffected by the building model uncertainties because these failures have larger effects on the designed FDD system than those of the modeling uncertainties. However, type-2 faults may have small effects on the FDD system and can be hidden as a consequence of modeling uncertainties. In any case, the theoretical framework developed in this paper can be extended to the synthesis of an FDD algorithm that is also robust to model uncertainties. Indeed, assume that the actual building dynamics is given by equation (7) where the nominal system matrices A and B are perturbed respectively by some matrices ΔA and ΔB of appropriate dimensions as a result of modeling uncertainties. Then, it is straightforward to represent the modeling uncertainties as some unmeasurable disturbances d_{Δ} entering the system through a distribution matrix E_{Δ} . It is noteworthy that the model equation (7) still holds for representing the uncertainties arising from modeling errors but now, with $E = [E_m \mid E_{um} \mid E_{\Delta}]$ and $d = [d_m^T \mid d_{um}^T \mid d_{\Delta}^T]^T$. Consequently, it is easily seen that the synthesis of the bilinear observer-based RAFE (9) remains unchanged provided that the set of unmeasurable disturbances is the set $\{d_{um}, d_{\Delta}\}$. The key proposition 3 for ensuring the robustness properties of the RAFE with regards to the unmeasured disturbances can be extended to this case with some mathematical technicalities. However, we will not dwell further into this aspect as the details depends strongly on the uncertainty modeling assumptions which need to be substantiated for determining the E_{Δ} -matrix. This is let for a future research. In this paper, the robustness against uncertainties arising from the mismatch between the model and the real building has been dealt in a practical way at the fault detection stage by setting judiciously the detection threshold.

B. Implementation of the robust fault diagnosis methodology

In this section, the implementation of the overall robust fault diagnosis methodology consisting of fault detection, isolation, and adaptive estimation of VAV damper faults is presented. The design matrices L, G, β_1 , and β_2 of the RAFE (9) are synthesized using Propositions 3 and 4. From Remark 1, faults in up to $n-1$ VAV dampers can be estimated using the RAFE (9). Therefore, prior to the estimation of a fault, it is required to identify which zones have a faulty VAV damper. Let us denote by RG the overall residual generator. This generator is composed of n residual generators $RG_i, i = 1, 2, \dots, n$, where each RG_i outputs a signal $r_i(t)$ which is a vector belonging to $\mathbb{R}^n$ and computed using the formula

$$r_i(t) = \Upsilon_i G \zeta_y(t, \hat{f}_a = 0) \quad (24)$$

with $\Upsilon_i \in \mathbb{R}^{n \times n}$ denoting a full rank matrix associated with the i -th residual signal of the bank. The signal $\zeta_y(t, \hat{f}_a = 0)$ in (24) is computed from (13) by substituting $\hat{f}_a = 0$. So the

overall residual signal of RG is

$$r(t) = [r_1^T(t) \mid r_2^T(t) \mid \cdots \mid r_n^T(t)]^T$$

$$= \left[\begin{array}{c|c|c|c} \begin{bmatrix} r_{11} \\ r_{12} \\ \vdots \\ r_{1n} \end{bmatrix}^T & \begin{bmatrix} r_{21} \\ r_{22} \\ \vdots \\ r_{2n} \end{bmatrix}^T & \cdots & \begin{bmatrix} r_{n1} \\ r_{n2} \\ \vdots \\ r_{nn} \end{bmatrix}^T \end{array} \right]^T$$

and the square of the Euclidean norm of $r(t)$ is given by $\|r(t)\|^2 = \sum_{i=1}^n \sum_{j=1}^n r_{ij}^2$. The matrix Υ_i in (24) is designed such that whenever there is a damper fault in the VAV terminal unit of the i -th zone, then only the i -th element of the vector $r_i(t)$ is zero and the i -th elements of the other residual vectors $r_k(t), \forall k \neq i$, are greater than some threshold ρ , that is,

$$\left(\begin{array}{l} (r_{ii}(t) = 0) \wedge \\ (|r_{ki}(t)| \geq \rho) \end{array} \right) \stackrel{\text{holds}}{:=} \begin{cases} \text{true,} & \text{fault is in } i\text{-th zone} \\ \text{false,} & i\text{-th zone is fault-free} \end{cases} \quad (25)$$

where $r_{ki}(t)$ denotes the i -th element of vector $r_k(t)$, $|\bullet|$ denotes the modulus and the symbol " $\wedge$ " denotes the conjunction operation. Using (25), single and multiple VAV damper faults can be successfully isolated. Fault isolation is preceded by a fault detection stage carried out as

$$\begin{aligned} \|r(t)\| &< \varrho, \quad \text{for no faults} \\ \|r(t)\| &\geq \varrho, \quad \text{for VAV damper faults} \end{aligned} \quad (26)$$

where ϱ is a given threshold value with units in $^\circ C$. The implementation of the complete robust fault detection and diagnosis methodology is shown in the block diagram of Fig. 2, where based on the measured signals, (z, u, d_m) , a fault occurrence is detected using the detection logic (26) followed by an isolation of the faulty VAV dampers using the rule (25), and finally the adaptive fault estimator is triggered for estimating precisely the magnitude of the fault using (9) with parameter matrices computed from Propositions 3 and 4. Note that two RAFEs are required for the implementation of the FDD methodology. In the event that VAV damper faults occur simultaneously in all the zones, then both the RAFEs are used for the estimation of the faults, with RAFE_1 used to estimate the damper faults occurred in $n - 1$ zones, while RAFE_2 is required for the estimation of the fault occurred in the n -th zone. If there occurs a single fault or multiple faults in up to $n - 1$ zones, then the faults estimates are obtained using only RAFE_1 .

IV. SIMULATION RESULTS

In order to demonstrate the effectiveness of the complete fault diagnosis methodology, a case study of a one-storey building containing three zones is simulated using the physics-based simulator SIMBAD, which is a Matlab/Simulink toolbox developed by CSTB [39] for prototyping and analysing control performance and monitoring methods in building HVAC systems. The SIMBAD simulator, being able to virtually reconstruct real buildings behaviors based on their actual configurations and different weather and occupancy conditions as well as use patterns, has been extensively used for testing and validating control and monitoring algorithms by industry

TABLE II
FAULT SCENARIOS FOR EVALUATION OF THE FAULT DIAGNOSIS METHODOLOGY

Scenario	Case	Fault Type	Faulty zones
1 (mutually exclusive)	1	Lock-in-place	1, 2, 3
	2	Bias	
2 (simultaneous)	1	Lock-in-place	1, 2, 3
	2	Bias	

TABLE III
FAULT SCENARIO - 1

Case	Fault Type	Damper-1	Damper-2	Damper-3
1	Lock-in-place (Airflow rate)	0.67 (0.061 kg/s)	0.68 (0.064 kg/s)	0.7 (0.071 kg/s)
2	Bias (Airflow rate)	-0.33 (0.061 kg/s)	-0.30 (0.071 kg/s)	-0.32 (0.064 kg/s)

engineers in preliminary steps before coding and deploying these algorithms in real building automation systems [40].

A. Setup of the controlled HVAC-VAV system

The building considered is a non-residential building consisting of three office spaces (zones), as illustrated in Fig. 3. The working hours are from 08h00 to 12h00 in the morning and from 14h00 to 18h00 in the afternoon for five days a week. The occupancy profile in Zone-3, Zone-2 and Zone-1 is set to five, four and six respectively, which contribute to the thermal load of the building. The minimum and maximum airflow rate possible by the VAV boxes of each zone corresponding to the damper position at 0 and 1 are $u_{min} = 0.49 \times 10^{-3} \text{ kg/s}$ and $u_{max} = 0.3163 \text{ kg/s}$, respectively. The three zones in this building have three external walls, denoted by $T_{extw,i}$, $i = 1, 2, 3$, which interacts with the ambient weather, and three internal walls, denoted with $T_{intw,12}$, $T_{intw,13}$ and $T_{intw,23}$, which interacts with neighboring zones. The numerical values of the building parameters A, B, D, E are taken from [37]. The supply temperature T_s for conditioning the air is kept at $27^\circ C$. Based on the occupancy profile and weather data, the thermal load in each zone, which acts as a disturbance in (Σ) , is computed within SIMBAD. The disturbance signals are shown in Fig. 4. The sensor noise in the measurement of zone temperature is simulated with zero mean and variance of 20×10^{-5} .

With reference to the closed-loop control of the multizone system shown in Fig. 1, the variable speed fan in the air-handling unit drives appropriate volume air mass to the input of VAV terminal units of the individual zones. The VAV boxes are pressure independent terminal units [41], which are typically controlled by a cascaded control strategy consisting of a master controller which is the temperature controller and a slave controller which controls the airflow rate. The outer-loop temperature control is a multivariable PI (Proportional + Integral) control law $u(t) = K_P(z(t) - y_{set}(t)) + K_I \int_0^t (z(\tau) - y_{set}(\tau)) d\tau$, which regulates each zone temperature to its reference value of $T_{z,set} = 20 \times 1.3^\circ C$, with the controller

Fig. 2. Robust fault detection and diagnosis methodology for multizone building VAV terminal units

Fig. 3. Envelope of the one-storey building

Fig. 4. Internal gain and outside ambient air temperature

gains selected as $K_P = I_3$ and $K_I = 0.0005I_3$. Under the event of VAV dampers faults, the desired volume of air is not delivered to the zone, thereby affecting the thermal comfort of occupants and the energy-efficient operation of the overall HVAC system.

B. Fault Diagnosis

We consider four different fault cases categorized in two different scenarios for demonstrating the effectiveness of the proposed fault diagnosis methodology, where the measured disturbance variable is Q_{int} and the unmeasured disturbance variable is T_{oa} . Nevertheless, there could be uncertainties in the measured disturbance variable also, since the mathematical bilinear model used in simulation scenarios captures the approximate dynamics of the overall HVAC building system simulated with SIMBAD toolbox, which is typically suitable for control and monitoring purposes. It is now apparent that uncertainties in the sense of the mismatch between the plant and the mathematical model are inherently present in simulations. It is shown through numerical simulations in subsequent sections that the developed fault estimator is not only robust to unmeasured disturbances as rigorously proved in the previous sections but also with regards to inherently unavoidable modelling uncertainties.

In Scenario-1, multiple VAV damper faults (lock-in-place and bias) occur in a mutually exclusive manner, and in Scenario-2, simultaneous faults occur in all zones. All considered fault cases are given in Table II. For both the scenarios, the design matrices for the robust adaptive fault estimator are selected as $L = 0.0005 \begin{bmatrix} I_3 \\ 0_{6 \times 3} \end{bmatrix}$, $\beta_1 = 0.002I_3$, $\beta_2 = 0.0005I_3$. For generating the residual signals for robust fault detection and isolation in the three-zones building, three matrices $\Upsilon_i, i =$

$$1, 2, 3 \text{ are selected as } \Upsilon_1 = V^{-1} \begin{bmatrix} 10 & 0 & 0 \\ 0 & \frac{20}{3} & -\frac{20}{\sqrt{3}} \\ 0 & 0 & 10 \end{bmatrix} V, \Upsilon_2 = V^{-1} \begin{bmatrix} 10 & 0 & 0 \\ 0 & 10 & 0 \\ \frac{20}{3} & 0 & -\frac{20}{\sqrt{3}} \end{bmatrix} V, \Upsilon_3 = V^{-1} \begin{bmatrix} \frac{10}{\sqrt{3}} & 10 & 0 \\ \frac{20}{3} & -\frac{20}{\sqrt{3}} & 0 \\ 0 & 0 & 10 \end{bmatrix} V,$$

where the columns of V are the generalized eigenvectors of the matrix G .

1) *Scenario 1: Mutually exclusive occurring multiple faults:*

In this scenario, multiple VAV damper faults occur in the building HVAC system. Two fault cases are considered, where in the first case a lock-in-place failure appears in VAV dampers of zone-3, zone-2, and zone-1 at the end of day 5, 3, and 1 respectively and disappears or is corrected in the respective zones by the end of the next day. The lock-in-places and their corresponding constant airflow rates are given in Table III. Under this fault, the zones temperature are shown in Fig. 5(a). It is worth noticing that the ‘intense’ variations in zone temperature measurement appear due to the departure and arrival of occupants. The norm of the residual signal is recorded in Fig. 5(b). Selecting the threshold values as $\rho = 2^\circ\text{C}$, and $\varrho = 5^\circ\text{C}$, the fault is successfully detected at the end of day 1, 3, and 5 according to (26). Moreover, when the damper fault disappears then the norm of the residual signal is again less than the threshold value. For fault isolation, the vector signals $r_i(t)$, $i = 1, 2, 3$, illustrated in Fig. 5(c), are used to determine the faulty damper. Recall that if the fault has occurred in the i -th zone, then the i -th signal of the vector $r_i(t)$ is zero while that of the other vectors, i.e. $r_{k \neq i}(t)$, satisfy the condition given in (25). This can be visualized in Fig. 5(c) that during day 2, the signal $r_{11}(t)$ is zero, and the signals $r_{21}(t)$, $r_{31}(t)$ satisfy $|r_{21}(t)| \geq 2$ and $|r_{31}(t)| \geq 2$. Also, during the same time, other signals in Fig. 5(c), satisfy $|r_{ik}| < 2, \forall i = 1, 2, 3$ and $k = 2, 3$, which implies that the fault has occurred only in zone-1. Therefore, using RAFE_1 , the fault in zone-1 is estimated, which is illustrated in Fig. 6. Now after day 2, the fault in zone-1 disappears and a fault in zone-2 appears at the end of day 3. From Fig. 5(b), the norm of the residual signal satisfies the condition given in (26). On checking the residual signals in Fig. 5(c), we have $r_{22}(t) = 0$ and $|r_{12}(t)| \geq 2, |r_{32}(t)| \geq 2$, which indicates that there is a fault in zone-2. Similarly, the lock-in-place failure of the VAV damper of zone-3 are detected and isolated. Based on the isolation of faulty dampers, only RAFE_1 is triggered in scenario-1, case-1 for the accurate estimation of the fault. The robust fault estimates under the occurrence of fault-type of case 1 from Table III are shown in Fig. 6 where it is established that the adaptive fault estimation filter driven by noisy sensor measurement and without having any information of d_{um} successfully reconstructs the airflow rates at the input of the zones in the fault-free mode as well as in the faulty condition.

For bias faults, a bias is considered in all VAV dampers occurring and disappearing as above. A constant bias in damper position is also given in Table III. Since the building HVAC system operates under feedback control, the controller output also saturates under the bias faults. The zones temperatures, the norm of the residual, and the signals generated by the bank of residuals are shown in Fig. 7. The output of the controller and the successful estimation of actual airflow rates are illustrated in Fig. 8. It is worth noting that the intense signature of measurement noise is no longer present on the fault estimates unlike during the fault-free operation.

TABLE IV
FAULT SCENARIO - 2

Case	Fault Type	Damper-1	Damper-2	Damper-3
1	Lock-in-place (Airflow rate)	0.7 (0.071 kg/s)	0.6 (0.041 kg/s)	0.6 (0.041 kg/s)
2	Bias (Airflow rate)	-0.33 (0.071 kg/s)	-0.32 (0.064 kg/s)	-0.35 (0.055 kg/s)

2) *Scenario 2: Simultaneously occurring multiple faults:*

In this scenario, similar types of faults as above occur in zone-1 and zone-2 simultaneously at the end of day 2, and in zone-3 at the end of day 4. Unlike above, both type of faults once occurred do not disappear. The damper positions and the corresponding airflow rates at the output of the VAV boxes are listed in Table IV. Under the lock-in-place failure, the temperature of the zones is shown in Fig. 9(a). The norm of the residual signal generated by the bank of residual generators is shown in Fig. 9(b), and the individual residual signals are recorded in Fig. 9(c). From Fig. 9(b), the normed signal exceeds the threshold value of 5°C as soon as the lock-in-place failure occurs in VAV dampers at the end of the day. According to the fault isolation logic (25), the signal $r_{11}(t)$ in Fig. 9(c) is zero, while the signals $r_{21}(t)$, $r_{31}(t)$ satisfy $|r_{21}(t)| \geq 2$ and $|r_{31}(t)| \geq 2$, which implies that a fault has occurred in zone-1. It is also noticed that after day 2, the signal $r_{22}(t)$ is zero, while the signals $r_{12}(t)$, $r_{32}(t)$ satisfy $|r_{12}(t)| \geq 2$ and $|r_{32}(t)| \geq 2$, which implies that a fault has also occurred in zone-2. At the same time, the signals $r_{i3}(t)$, $i = 1, 2, 3$ satisfy $|r_{i3}(t)| < 2$ until the end of day 4. However, the signals $r_{13}(t) \geq 2$ and $r_{23}(t) \geq 2$ with $r_{33}(t) = 0$ after that day. This implies that now a fault has also occurred in zone-3 after day 4. Finally, since the VAV damper fault is now present in all the zones after day 4, only RAFE_1 is activated up to day 4, while the estimator RAFE_2 is activated afterwards. The robust adaptive estimation of the magnitude of the VAV damper lock-in-place failures is illustrated in Fig. 10. Similarly, the signals corresponding to bias faults occurring in all zones are recorded in Fig. 11 and 12.

V. CONCLUSION

In this work, a new robust fault diagnosis methodology is developed for multizone building VAV terminal units which ensures some robustness properties with regards to unmeasured disturbances and sensor noises. The methodology consists of detecting, isolating and adaptively estimating any single or multiple VAV damper failures. In the considered context, the unmeasured disturbance is the outdoor temperature which causes the thermal dynamics of the building to vary significantly over a large operating region, thus making accurate fault diagnosis a challenging task due to the non-linearity underlying this wide operating region. Unlike the usual building dynamics linearization-based approaches, the robust FDD algorithm is built upon a bilinear observer derived from the real thermal dynamics of the building over its entire

Fig. 5. Scenario-1, Case-1

Fig. 6. Scenario-1, Case-1: Output signals of the PI controller and that of the VAV boxes

operating region essentially determined by the outdoor climate conditions, and in particular the outside temperature. The key feature of this observer is that it is augmented by a judiciously designed filter-based adjustment law for adaptive fault

estimation that achieves a reduced sensitivity to measurement noises and a decoupling of the estimates from the unmeasured disturbance. Exhaustive simulation results using the Matlab-based simulator SIMBAD are reported to demonstrate the effectiveness of the robust FDD methodology under different VAV damper fault scenarios in a one-story building with three zones. Though the present work was mainly motivated by dealing with the outdoor temperature as an unmeasured disturbance, a byproduct of the developed methodology is that it might also be applicable to precisely estimate the faults under the unavailability of internal heat gains from a zone of the building as well as under modeling uncertainties.

REFERENCES

- [1] J. Yamé, T. Jain, and D. Sauter, "An online controller redesign based fault-tolerant strategy for thermal comfort in a multi-zone building," in *IEEE Conference on Control Applications*, Sydney, NSW, Australia, 2015, pp. 1901–1906.
- [2] M. Anderson, M. Buehner, P. Young, D. Hittle, C. Anderson, J. Tu, and D. Hodgson, "MIMO robust control for HVAC systems," *IEEE Transactions on Control Systems Technology*, vol. 16, no. 3, pp. 475–483, 2008.
- [3] Y. Ma, J. Matuško, and F. Borrelli, "Stochastic model predictive control for building HVAC systems: Complexity and conservatism," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 1, pp. 101–116, 2014.
- [4] X. Qin, S. Lysecky, and J. Sprinkle, "A data-driven linear approximation of HVAC utilization for predictive control and optimization," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 2, pp. 778–786, 2014.
- [5] M. Najafi, D. M. Auslander, P. L. Bartlett, P. Haves, and M. D. Sohn, "Application of machine learning in the fault diagnostics of air handling units," *Applied Energy*, vol. 96, pp. 347 – 358, 2012.

Fig. 7. Scenario-1, Case-2

Fig. 8. Scenario-1, Case-2: Output signals of the PI controller and that of the VAV boxes

- [6] S. Lee and F. Yik, "A study on the energy penalty of various air-side system faults in buildings," *Energy and Buildings*, vol. 42, no. 1, pp. 2 – 10, 2010.
- [7] B. Gunay, W. Shen, and C. Yang, "Characterization of a building's operation using automation data: A review and case study," *Building and Environment*, vol. 118, pp. 196 – 210, 2017.

- [8] A. Montazeri and S. M. Kargar, "Fault detection and diagnosis in air handling using data-driven methods," *Journal of Building Engineering*, vol. 31, p. 101388, 2020.
- [9] B. Sun, P. B. Luh, Q. Jia, Z. O'Neill, and F. Song, "Building Energy Doctors: An SPC and Kalman Filter-Based Method for System-Level Fault Detection in HVAC Systems," *IEEE Transactions on Automation Science and Engineering*, vol. 11, no. 1, pp. 215–229, Jan 2014.
- [10] D. Li, Y. Zhou, G. Hu, and C. J. Spanos, "Identifying unseen faults for smart buildings by incorporating expert knowledge with data," *IEEE Transactions on Automation Science and Engineering*, vol. 16, no. 3, pp. 1412–1425, 2018.
- [11] S. M. Namburu, M. S. Azam, J. Luo, K. Choi, and K. R. Pattipati, "Data-driven modeling, fault diagnosis and optimal sensor selection for HVAC chillers," *IEEE Transactions on Automation Science and Engineering*, vol. 4, no. 3, pp. 469–473, July 2007.
- [12] A. Glass, P. Gruber, M. Roos, and J. Todtli, "Qualitative model-based fault detection in air-handling units," *IEEE Control Systems Magazine*, vol. 15, no. 4, pp. 11–22, 1995.
- [13] 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 *IEEE International Conference on Control Applications (CCA)*, Denver, CO, USA, Sept 2011, pp. 70–75.
- [14] H. Shahnazari, P. Mhaskar, J. M. House, and T. I. Salsbury, "Heating, ventilation and air conditioning systems: Fault detection and isolation and safe parking," *Computers and Chemical Engineering*, vol. 108, pp. 139–151, 2018.
- [15] 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.
- [16] T. Darure, J. J. Yamé, and F. Hamelin, "Fault-adaptive control of VAV damper stuck in a multizone building," in *3rd Conference on Control and Fault-Tolerant Systems (SysTol)*. IEEE, 2016, pp. 170–176.
- [17] V. Reppa, P. Papadopoulos, M. M. Polycarpou, and C. G. Panayiotou, "A distributed architecture for HVAC sensor fault detection and isolation," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 4, pp. 1323–1337, 2014.

Fig. 9. Scenario-2, Case-1

Fig. 10. Scenario-2, Case-1: Output signals of the PI controller and that of the VAV boxes

- [18] M. S. A. T. Jain, and J. J. Yamé, "Bilinear model-based diagnosis of lock-in-place failures of variable-air-volume HVAC systems of multi-zone buildings," *Journal of Building Engineering*, vol. 28, p. 101023, 2020.
- [19] H. Wang, Y. Chen, C. W. Chan, and J. Qin, "A robust fault detection

- and diagnosis strategy for pressure-independent VAV terminals of real office buildings," *Energy and Buildings*, vol. 43, no. 7, pp. 1774 – 1783, 2011.
- [20] 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.
- [21] 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.
- [22] J. Liang and R. Du., "Model-based fault detection and diagnosis of HVAC systems using support vector machine method," *International Journal of Refrigeration*, vol. 30, pp. 1104–1114, 2007.
- [23] Y. Zhao, T. Li, X. Zhang, and C. Zhang, "Artificial intelligence-based fault detection and diagnosis methods for building energy systems: Advantages, challenges and the future," *Renewable and Sustainable Energy Reviews*, vol. 109, pp. 85 – 101, 2019.
- [24] T. Jain, J. Yamé, and D. Sauter, *Active Fault-Tolerant Control Systems: A Behavioral System Theoretic Perspective*, 1st ed., ser. Studies in Systems, Decision and Control. Springer International Publishing, 2018.
- [25] H. Shahnazari and P. Mhaskar, "Actuator and sensor fault detection and isolation for nonlinear systems subject to uncertainty," *International Journal of Robust and Nonlinear Control*, vol. 28, no. 6, pp. 1996–2013, 2018.
- [26] M. Du, J. Scott, and P. Mhaskar, "Actuator and sensor fault isolation of nonlinear process systems," *Chemical Engineering Science*, vol. 104, pp. 294–303, 2013.
- [27] F. Oldewurtel, A. Parisio, C. N. Jones, D. Gyalistras, M. Gwerder, V. Stauch, B. Lehmann, and M. Morari, "Use of model predictive control and weather forecasts for energy efficient building climate control," *Energy and Buildings*, vol. 45, pp. 15–27, 2012.
- [28] D. Kim, J. Cai, K. B. Ariyur, and J. E. Braun, "System identification for building thermal systems under the presence of unmeasured disturbances in closed loop operation: Lumped disturbance modeling approach," *Building and Environment*, vol. 107, pp. 169–180, 2016.
- [29] B. Dong and K. P. Lam, "A real-time model predictive control for building heating and cooling systems based on the occupancy behavior

Fig. 11. Scenario-2, Case-2

Fig. 12. Scenario-2, Case-2: Output signals of the PI controller and that of the VAV boxes

- pattern detection and local weather forecasting,” in *Building Simulation*, vol. 7, no. 1. Springer, 2014, pp. 89–106.
- [30] J. R. Dobbs and B. M. Hency, “Model predictive hvac control with online occupancy model,” *Energy and Buildings*, vol. 82, pp. 675–684, 2014.

- [31] S. C. Bengea, A. D. Kelman, F. Borrelli, R. Taylor, and S. Narayanan, “Implementation of model predictive control for an hvac system in a mid-size commercial building,” *HVAC&R Research*, vol. 20, no. 1, pp. 121–135, 2014.
- [32] S. Goyal, P. Barooah, and T. Middelkoop, “Experimental study of occupancy-based control of hvac zones,” *Applied Energy*, vol. 140, pp. 75–84, 2015.
- [33] A. R. Coffman and P. Barooah, “Simultaneous identification of dynamic model and occupant-induced disturbance for commercial buildings,” *Building and Environment*, vol. 128, pp. 153–160, 2018.
- [34] D. Lazos, A. B. Sproul, and M. Kay, “Optimisation of energy management in commercial buildings with weather forecasting inputs: A review,” *Renewable and Sustainable Energy Reviews*, vol. 39, pp. 587–603, 2014.
- [35] Z. Gao, X. Liu, and M. Chen, “Unknown input observer-based robust fault estimation for systems corrupted by partially decoupled disturbances,” *IEEE Transactions on Industrial Electronics*, vol. 63, no. 4, pp. 2537–2547, 2016.
- [36] A. J. Krener and A. Isidori, “Linearization by output injection and nonlinear observers,” *Systems & Control Letters*, vol. 3, no. 1, pp. 47–52, 1983.
- [37] M. S. A. T. Jain, and J. J. Yamé, “Output Injected Nonlinear Observer for Diagnosing Faults in Multi-Zone Building,” in *2019 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, Hong Kong, China, July 2019, pp. 660–665.
- [38] C. Chen, *Linear system theory and design*. Oxford University Press, Inc, 1995.
- [39] SIMBAD, “SIMulator of Building And Devices,” Centre Scientifique et Technique du Batiment, France, 2005.
- [40] A. Husaunndee, H. Vaezi-Nejad, and J. Visier, “Use of simulation tools for design and test of control and energy management systems in intelligent buildings,” *International journal of solar energy*, vol. 21, no. 2-3, pp. 111–130, 2001.
- [41] G. Levermore, *Building Energy Management Systems. Applications to Low-Energy HVAC and Natural Ventilation Control, 2nd Ed.* E & FN Spon, Taylor & Francis Group, 2000.