Partial-Distributed Filtering for Fault Detection, Isolation and Accommodation in Natural-Gas Pipelines

Khadija Shaheen, *Student Member, IEEE,* Apoorva Chawla, *Member, IEEE,* Ferdinand Evert Uilhoorn, and Pierluigi Salvo Rossi, *Senior Member, IEEE*

***Abstract*—This paper explores an innovative method for distributed state estimation aimed at reducing computational complexity while detecting sensor faults in natural gas pipelines. The proposed framework utilizes a partial-distributed ensemble Kalman filter (EnKF), comprising linear local filters and a non-linear main filter. The main filter handles non-linear computations during the time update, while the simultaneous operation of linear local filters manages linear computations during the measurement update. These local filters generate distinct local state estimates based on their specific sensor measurements, which are then transmitted to an information mixer to compute fault-free state estimates. Moreover, a fault diagnosis strategy is developed using local state variances and residuals. Faulty sensors are identified and isolated by comparing these metrics against a threshold. Additionally, an adaptive thresholding approach is incorporated to enhance effective fault identification. The effectiveness of the proposed technique is demonstrated in systems characterized by high nonlinearity and dimensionality, and featuring simultaneous multiple sensor faults, through extensive simulations and comparative analyses.**

***Index Terms*—Partial-distributed ensemble Kalman filter, model-based technique, state estimation, fault diagnosis, adaptive threshold, multiple sensor faults, natural-gas pipelines.**

I. Introduction

TIMELY monitoring of natural gas pipelines is crucial for ensuring the overall safety and reliability of the system. As digital technologies advance, numerous monitoring systems have emerged for urban gas pipelines to detect potential leaks [1], [2]. These systems employ several gas-monitoring sensors to continuously observe gas pipelines, intending to identify potential leaks and enhance the reliability of pipeline infrastructures [3], [4], [5]. However, sensors installed in the pipelines may occasionally malfunction, resulting in delayed or overlooked leak detections, potentially causing severe failures [6]. Therefore, there is a critical need to develop effective techniques for timely identifying sensor faults within gas-monitoring systems.

Sensor fault diagnosis techniques can be broadly categorized into two groups: data-driven [7] and model-based approaches [8]. Data-driven methods utilize real-time or historical data for sensor fault diagnosis, relying on extensive databases that render them impractical for real-time applications [9], [10]. In contrast, model-based methods solely depend on an accurate system model. These approaches identify faults by comparing actual process measurements with model-predicted values [11]. Among model-based methods, the Kalman filter (KF) and its variants are widely preferred for their robustness and effectiveness in optimal state estimation [12], [13]. In the context of large distributed systems, there are two prevalent architectures for the KF: centralized and distributed. In the centralized architecture, a single central filtering unit processes the sensor data. Further, this approach possesses high accuracy while offering substantial computations for large-scale systems with numerous sensors [13]–[15]. On the other hand, the distributed architectures employ multiple local filters that work in parallel, with their outputs combined through a master filter. This approach reduces overall computations and improves decision-making capabilities [16]–[19].

In [20], the sensor and process fault diagnosis is investigated using both distributed and centralized multi-sensor data fusion architectures, incorporating adaptive extended Kalman filter (EKF) techniques. The centralized architecture ensures high estimation accuracy using a single filter, however, it is less robust against sensor faults. Conversely, the distributed architecture, employing distinct local filters for each sensor, faces challenges in handling nonlinearities and multiple sensor faults. In [21], a sensor-fusion strategy is presented, employing multiple KFs customized for specific defects in nonlinear systems. However, this approach is constrained by significant computational costs. In [22], a sensor-fusion technique based on the unscented Kalman filter (UKF) is investigated for monitoring a gas turbine engine, involving four different combinations of local filters for the installed sensors. However, this method encounters difficulties in effectively identifying multiple sensor faults. Moreover, [23] developed a UKF-based distributed sensor-fusion system for microgrid applications. Nevertheless, it is primarily designed for single-fault scenarios and becomes computationally intensive when dealing with

This work was partially supported by the Research Council of Norway under the project SIGNIFY within the IKTPLUSS framework (project nr. 311902).

K. Shaheen and A. Chawla are with the Department of Electronic Systems, Norwegian University of Science and Technology, 7491 Trondheim, Norway(e-mail: shaheen.khadija@ntnu.no; apoorva.chawla@ntnu.no).

F. E. Uilhoorn is with the Department of Gas Engineering, Warsaw University of Technology, Warsaw, Poland (e-mail: ferdinand.uilhoorn@pw.edu.pl).

P. Salvo Rossi is with the Department of Electronic Systems, Norwegian University of Science and Technology, 7491 Trondheim, Norway, and with the Department of Gas Technology, SINTEF Energy Research, 7491 Trondheim, Norway (e-mail: salvorossi@ieee.org).

multiple sensor faults. In addition, [24], [25] address the fault diagnosis using distributed filtering in natural gas pipelines, with a primary focus on addressing single-sensor faults. Furthermore, the work in [26] utilizes an EKF-based partially distributed architecture to detect single sensor and process faults in a gas turbine engine. However, it is not well-suited for addressing multiple sensor faults in highly nonlinear and large-scale systems. Moreover, [27] addresses the issue of simultaneous multiple sensor faults in natural gas pipelines using an ensemble Kalman filter (EnKF)-based distributed architecture. However, this approach faces substantial computational costs due to its fully distributed structure.

The existing distributed fault diagnosis methods present a notable challenge due to their elevated computational complexity, particularly in the case of large-scale systems encompassing numerous subsystems or local filters. In these systems, each subsystem typically requires repetitive nonlinear calculations associated with the system model, leading to a substantial increase in overall computational burden. To address these challenges, it is essential to develop new distributed architectures that reduce redundant nonlinear computations and effectively handle multiple sensor faults while maintaining low computational complexity.

In this paper, we design an innovative partial-distributed architecture based on EnKF, aiming to address the computational challenges associated with nonlinear state estimation and enable effective diagnosis of multiple sensor faults in natural gas pipelines. The proposed framework separates the nonlinear computations from the local filters and delegates them to the main filter. The main filter handles the time update and information fusion, while the local filters focus on the measurement updates, resulting in a significant reduction in the overall system's computational complexity and improving the fault diagnosis capabilities. The EnKF-based architecture provides specific advantages, particularly in dealing with high-dimensional and non-linear system of natural-gas pipeline [28]. It also enhances computational efficiency and overall decision-making compared to the conventional distributed architecture. Additionally, a novel fault-tolerant approach based on local state variance and residual is introduced. This approach monitors these metrics against a threshold for fault identification and facilitates the isolation of faulty state estimates from the information fusion process. To further improve detection accuracy, an adaptive thresholding technique is also incorporated that can dynamically adjust the threshold value based on evolving system dynamics. The final global estimate is obtained by combining all the non-faulty local filter estimates using the information fusion.

The structure of the paper is as follows: Section II elaborates on the system model of natural gas pipelines; the development of the sensor fault detection, isolation and accommodation (SFDIA) architecture is detailed in Section III; the presentation and discussion of the achieved performance are covered in Section IV; and Section V provides concluding remarks.

II. Natural Gas Transient Flow Model

The system model characterizing the transient flow of natural gas in pipelines can be expressed by a set of hyperbolic partial differential equations (PDEs) [29], given as

$$\frac{\partial \boldsymbol{x}}{\partial t} + \boldsymbol{A}(\boldsymbol{x})\frac{\partial \boldsymbol{x}}{\partial s} + \boldsymbol{\zeta}(\boldsymbol{x}) = 0 \ , \tag{1}$$

where the symbol s and t represent space and time, respectively. The spatial and temporal dynamics can be collectively defined as $\Omega = \{(s,t) : 0 \leq s \leq L, 0 \leq t \leq t_f\}$, where L and t_f denote the pipeline length and the time period, respectively. Additionally, the vector $\boldsymbol{x} = [p, \dot{m}, T]^T$ represents the state vector, where p, $\dot{m}$ and T denote pressure, flow rate and temperature, respectively, and the matrix $\boldsymbol{A}(\boldsymbol{x}) \in \mathbb{R}^{3\times 3}$ signifies the coefficient matrix, given by Eq. (2) [30], [31]. Further, the vector $\boldsymbol{\zeta}(\boldsymbol{x}) \in \mathbb{R}^{3\times 1}$ can be defined as

$$\boldsymbol{\zeta}(\boldsymbol{x}) = \begin{bmatrix} -\frac{a_s^2\alpha_1(Aqp+RT\dot{m}wz)}{A^2TC_pp} & w & -\frac{a_s^2\alpha_2(Aqp+RT\dot{m}wz)}{A^2C_pp^2} \end{bmatrix}^T \ ,$$

where $\alpha_1 = 1 + \frac{T}{z}\left(\frac{\partial z}{\partial T}\right)_p$, $\alpha_2 = 1 - \frac{p}{z}\left(\frac{\partial z}{\partial p}\right)_T$. The symbols a_s z, R, C_p, A, q, w, and ρ represent the isentropic wave speed, gas compressibility factor, ideal gas constant, specific heat at constant pressure, cross-sectional area, heat flow per unit length and time, frictional force per unit length of the pipe and density, respectively [32]. Furthermore, the transport and thermodynamic properties, specifically z and C_p, are obtained using the GERG-2004 [33].

Further, the numerical method of lines is applied to solve the transient model, utilizing spatial discretization to convert the system of PDEs into a set of ordinary differential equations (ODEs) [34]. A 5-point, fourth-order finite difference method that introduces an error of order $\mathcal{O}(\Delta s^4)$, where Δs denotes the spatial step size, is employed to convert the system of PDEs in (1). The system of ODEs obtained after spatial discretization can be expressed as

$$\frac{d\boldsymbol{x}(t)}{dt} = \boldsymbol{B}(\boldsymbol{x})\boldsymbol{D}\boldsymbol{x}(t) - \boldsymbol{\Gamma}(\boldsymbol{x},t) \triangleq \boldsymbol{\varphi}(t,\boldsymbol{x}(t)) \ , \tag{3}$$

where $\boldsymbol{B}(\boldsymbol{x}) \in \mathbb{R}^{3n\times 3n}$ and $\boldsymbol{\Gamma}(\boldsymbol{x},t) \in \mathbb{R}^{3n\times 1}$ denote the assembled matrix of $\boldsymbol{A}(\boldsymbol{x})$ and the assembled column vector of $\boldsymbol{\zeta}(\boldsymbol{x})$, respectively. The computational matrix $\boldsymbol{D}$ is given in [29] and the modified state vector $\boldsymbol{x}(t)$ can be expressed as $\boldsymbol{x}(t) = [p_1(t), \ldots, p_i(t), \ldots, p_n(t), \dot{m}_1(t), \ldots, \dot{m}_i(t), \ldots, \dot{m}_n(t), T_1(t), , \ldots, T_i(t), \ldots, T_n(t)]^T$. The system of ODEs can be further converted into a state-space model using an effective numerical technique known as the fourth-order Runge-Kutta method [27]. When discretized using a constant time step, the progression of the solution over time can be described as

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \frac{\Delta t}{6}\left(\boldsymbol{k}_1 + 2\boldsymbol{k}_2 + 2\boldsymbol{k}_3 + \boldsymbol{k}_4\right) \ ,$$

$$\boldsymbol{A}(\boldsymbol{x}) = \begin{bmatrix} A - \frac{-\frac{\dot{m}(a_s^2\alpha_2 - RTz)}{Ap}}{\frac{a_s^2\alpha_2^2 C_p \dot{m}^2 - Ra_s^2\alpha_1^2\alpha_2\dot{m}^2 z}{AC_p p^2}} & \frac{\frac{a_s^2}{A}}{\frac{\dot{m}(\alpha_2 C_p a_s^2 - Rza_s^2\alpha_1^2 + RTC_p z)}{AC_p p}} & \frac{\frac{a_s^2\dot{m}\alpha_1}{AT}}{\frac{a_s^2\alpha_1\dot{m}^2(\alpha_2 C_p - R\alpha_1^2 z)}{ATC_p p}} \\ -\frac{RTa_s^2\alpha_1\alpha_2\dot{m}z}{AC_p p^2} & \frac{RTa_s^2\alpha_1 z}{AC_p p} & \frac{R\dot{m}z(a_s^2\alpha_1^2 + TC_p)}{AC_p p} \end{bmatrix} \tag{2}$$

where

$$\begin{aligned} \boldsymbol{k}_1 &= \boldsymbol{\varphi}(\boldsymbol{x}_k, t_k) \ , \\ \boldsymbol{k}_2 &= \boldsymbol{\varphi}\left(\boldsymbol{x}_k + \frac{\Delta t}{2}\boldsymbol{k}_1, t_k + \frac{\Delta t}{2}\right) \ , \\ \boldsymbol{k}_3 &= \boldsymbol{\varphi}\left(\boldsymbol{x}_k + \frac{\Delta t}{2}\boldsymbol{k}_2, t_k + \frac{\Delta t}{2}\right) \ , \\ \boldsymbol{k}_4 &= \boldsymbol{\varphi}\left(\boldsymbol{x}_k + \Delta t\boldsymbol{k}_3, t_k + \Delta t\right) \ , \end{aligned}$$

$\boldsymbol{x}_k = \boldsymbol{x}(t_k)$ and $t_k = k(\Delta t)$. Further, it satisfies the Courant-Friedrichs-Lewy condition, given in [29], [35], for numerical stability. The proposed partially-distributed EnKF based SFDIA architecture is discussed in the following section.

III. Proposed Architecture

Fig. 1 illustrates various steps of the proposed architecture. The sensor measurements are initially grouped into several unique sets of sensor measurements and are assigned to specific local filters. The proposed partial-distributed architecture subsequently performs fault-free state estimation using the following steps: (1) The main filter first carries out the nonlinear computation (time update). (2) The linear local filters then execute the measurement updates, estimating the local state vectors and their associated covariance matrices using their specific subset of measurements. (3) Then, sensor faults are identified by comparing the state variance vector with an adaptable threshold, computed using the state variance vector. (4) The faults are accommodated using the local state residual vector, and the faulty local state estimate is substituted with a corrected non-faulty *a priori* estimate. (5) The global estimates are finally computed using the corrected local state estimates in the information mixture. The final estimates are fed to the main filter for initialization during the subsequent iteration. In the following subsections, the state estimation and fault diagnosis by the proposed partial-distributed filter are discussed in detail.

A. Sensors grouping

Initially, the locals and their associated subset of measurements are defined, ensuring each local filter has unique measurements. With M sensors and N filters, each filter processes M/N measurements[1]. This grouping guarantees that the local filters generate similar state vector estimates using their specific measurements in the absence of faults, facilitating fault detection. This approach effectively identifies faulty sensors, thereby offering a resilient solution. However, it should be noted that this method is optimal only for uncorrelated sensor measurements. Subsequently, the design of the partial-distributed EnKF-based filter is discussed next.

[1] N/M is an integer number.

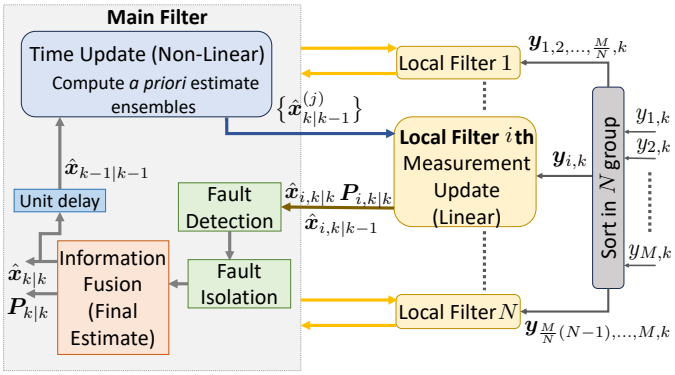


Fig. 1: The proposed partial-distributed architecture.

B. Partially-Distributed EnKF-based Architecture

The nonlinear state-space model can be expressed as

$$\begin{aligned} \boldsymbol{x}_k &= \boldsymbol{f}(\boldsymbol{x}_{k-1}, \boldsymbol{u}_{k-1}) + \boldsymbol{v}_k \ , \\ \boldsymbol{y}_{i,k} &= \boldsymbol{h}_i(\boldsymbol{x}_k, \boldsymbol{u}_k) + \boldsymbol{n}_{i,k} \ , \end{aligned}$$

where $\boldsymbol{f}(\cdot,\cdot) : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_x}$ is the nonlinear mapping of the main filter with $\boldsymbol{x}_k \in \mathbb{R}^{n_x \times 1}$ being the state vector during the kth time instant. Further, $\boldsymbol{y}_{i,k} \in \mathbb{R}^{n_y \times 1}$ denotes the measurement vector corresponding to the ith local filter, where $i = 1, 2, \ldots, N$. The input vector $\boldsymbol{u}_{k-1} \in \mathbb{R}^{n_u \times 1}$ includes both initial $\boldsymbol{u}_{in}$ and boundary $\boldsymbol{u}_{bc,k-1}$ conditions. Additionally, the mapping corresponding to the measurement model for the ith local filter is considered linear, i.e., $\boldsymbol{h}_i(\boldsymbol{x}_k, \boldsymbol{u}_k) = \boldsymbol{H}_i\boldsymbol{x}_k$. The aim is to reduce the dimension of the state vector through one-to-one mapping based on measurements allocated to each local filter. Moreover, $\boldsymbol{v}_k \in \mathbb{R}^{n_x \times 1} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_k)$ and $\boldsymbol{n}_{i,k} \in \mathbb{R}^{n_y \times 1} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_{i,k})$ represent the process noise and measurement noise, respectively.

The different steps of the state estimation procedure involved in the proposed partially distributed EnKF architecture are elaborated next.

Step 0: Initialization:

The initial state vector estimate $\hat{\boldsymbol{x}}_{0|0}$ is fixed in the first step.

Step 1: Main filter nonlinear computation:

An ensemble of samples, denoted by $\{\hat{\boldsymbol{x}}^{(j)}_{k-1|k-1}, 1 \leq j \leq N_e\}$, is generated to represent the distribution $p(\boldsymbol{x}_{k-1}|\mathbb{Y}_{k-1})$, with N_e being the ensemble size and $\mathbb{Y}_{k-1} = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_{k-1}\}$. The ensemble of the state vector estimate, denoted by $\mathcal{X}_{k-1|k-1} \in \mathbb{R}^{n_x \times N_e}$, is generated as

$$\hat{\boldsymbol{x}}^{(j)}_{k-1|k-1} = \hat{\boldsymbol{x}}_{k-1|k-1} + \boldsymbol{v}^{(j)}_{k-1} \ , \quad j = 1, 2, \ldots, N_e \ . \tag{4}$$

Further, the *a priori* ensemble associated with $p(\boldsymbol{x}_k|\mathbb{Y}_{k-1})$ can be obtained as

$$\hat{\boldsymbol{x}}_{k|k-1}^{(j)} = \boldsymbol{f}\left(\hat{\boldsymbol{x}}_{k-1|k-1}^{(j)}, \boldsymbol{u}_{k-1}^{(j)}\right) + \boldsymbol{v}_k^{(j)} \ . \tag{5}$$

This *a priori* ensemble is subsequently shared with N linear local filters, each incorporating its distinct local measurement subset, to generate *a posteriori* local state estimate $\hat{\boldsymbol{x}}_{i,k|k}$.

Step 2: Local filter linear computation:

In each local filter, an ensemble of measurement samples is generated. For the ith local filter, an ensemble $\left\{\hat{\boldsymbol{y}}_{i,k|k-1}^{(j)}, 1 \leq j \leq N_e\right\}$, representing $p(\boldsymbol{y}_{i,k-1}|\mathbb{Y}_{i,k-1})$, is generated, which can be obtained as

$$\hat{\boldsymbol{y}}_{i,k|k-1}^{(j)} = \boldsymbol{H}_i \hat{\boldsymbol{x}}_{k|k-1}^{(j)} + \boldsymbol{n}_{i,k}^{(j)}, \tag{6}$$

where $\mathbb{Y}_{i,k-1} = \{\boldsymbol{y}_{i,1}, \boldsymbol{y}_{i,2}, \ldots, \boldsymbol{y}_{i,k-1}\}$ and $\boldsymbol{y}_{i,k-1} = \left[y_{[\frac{M}{N}(i-1)+1],k-1}, \ldots, y_{[\frac{M}{N}(i-1)+\frac{M}{N}],k-1}\right]^T$. The mean and covariance of the measurement ensemble $\left\{\hat{\boldsymbol{y}}_{i,k|k-1}^{(j)}, 1 \leq j \leq N_e\right\}$ can be expressed as $\hat{\boldsymbol{y}}_{i,k|k-1} = \frac{1}{N_e}\sum_{j=1}^{N_e} \hat{\boldsymbol{y}}_{i,k|k-1}^{(j)}$ and $\boldsymbol{P}_{i,k|k-1}^{y} = \frac{1}{N_e-1}\boldsymbol{E}_{i,k|k-1}^{y}\left(\boldsymbol{E}_{i,k|k-1}^{y}\right)^T$, respectively, where $\boldsymbol{E}_{i,k|k-1}^{y} = \left[\left(\hat{\boldsymbol{y}}_{i,k|k-1}^{(1)} - \hat{\boldsymbol{y}}_{i,k|k-1}\right), \ldots, \left(\hat{\boldsymbol{y}}_{i,k|k-1}^{(N_e)} - \hat{\boldsymbol{y}}_{i,k|k-1}\right)\right]$. Likewise, the mean and covariance of the *a priori* state ensemble can be computed as $\hat{\boldsymbol{x}}_{i,k|k-1} = \frac{1}{N_e}\sum_{j=1}^{N_e} \hat{\boldsymbol{x}}_{i,k|k-1}^{(j)}$ and $\boldsymbol{P}_{i,k|k-1}^{x} = \frac{1}{N_e-1}\boldsymbol{E}_{i,k|k-1}^{x}\left(\boldsymbol{E}_{i,k|k-1}^{x}\right)^T$, respectively, where $\boldsymbol{E}_{i,k|k-1}^{x} = \left[\left(\hat{\boldsymbol{x}}_{i,k|k-1}^{(1)} - \hat{\boldsymbol{x}}_{i,k|k-1}\right), \ldots, \left(\hat{\boldsymbol{x}}_{i,k|k-1}^{(N_e)} - \hat{\boldsymbol{x}}_{i,k|k-1}\right)\right]$. The cross-covariance between $\boldsymbol{x}_{i,k|k-1}$ and $\boldsymbol{y}_{i,k|k-1}$, given $\mathbb{Y}_{i,k-1}$, can be evaluated as $\boldsymbol{P}_{i,k|k-1}^{xy} = \frac{1}{N_e-1}\boldsymbol{E}_{i,k|k-1}^{x}\left(\boldsymbol{E}_{i,k|k-1}^{y}\right)^T$.

Then, the *a priori* ensemble $\left\{\hat{\boldsymbol{x}}_{i,k|k-1}^{(j)}\right\}$ can be updated using the latest measurement $\boldsymbol{y}_{i,k}$ as

$$\hat{\boldsymbol{x}}_{i,k|k}^{(j)} = \hat{\boldsymbol{x}}_{i,k|k-1}^{(j)} + \boldsymbol{K}_{i,k}\left(\boldsymbol{y}_{i,k} - \hat{\boldsymbol{y}}_{i,k|k-1}^{(j)}\right), \tag{7}$$

$$\boldsymbol{K}_{i,k} = \boldsymbol{P}_{i,k|k-1}^{xy}\left(\boldsymbol{P}_{i,k|k-1}^{y}\right)^{-1}. \tag{8}$$

Finally, the *a posteriori* state estimate and covariance can be computed as

$$\hat{\boldsymbol{x}}_{i,k|k} = \frac{1}{N_e}\sum_{j=1}^{N_e} \hat{\boldsymbol{x}}_{i,k|k}^{(j)}, \boldsymbol{P}_{i,k|k} = \frac{1}{N_e-1}\boldsymbol{E}_{i,k|k}^{x}\left(\boldsymbol{E}_{i,k|k}^{x}\right)^T,$$

with $\boldsymbol{E}_{i,k|k}^{x} = \left[\left(\hat{\boldsymbol{x}}_{i,k|k}^{(1)} - \hat{\boldsymbol{x}}_{i,k|k}\right), \ldots, \left(\hat{\boldsymbol{x}}_{i,k|k}^{(N_e)} - \hat{\boldsymbol{x}}_{i,k|k}\right)\right]$. Following the state estimation using the sensor measurements, fault detection and isolation are performed.

Step 3: Fault detection and isolation:

The metric such as the state-variance vector, denoted by $\boldsymbol{\xi}_k \in \mathbb{R}^{n_x \times 1}$, is used for fault detection, which is defined as

$$\xi_k^{(\ell)} = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{x}_{i,k|k}^{(\ell)} - \frac{1}{N}\sum_{i=1}^{N}\left(\hat{x}_{i,k|k}^{(\ell)}\right)\right)^2, \tag{9}$$

where $\hat{x}_{i,k|k}^{(\ell)}$ specifies the ℓth element of the ith local state vector estimate $\hat{\boldsymbol{x}}_{i,k|k}$. Utilizing the grouping outlined in Subsection IIIA, each local subset of measurements only affects its corresponding elements in the local state vector [27]. Thus, this grouping enables the generation of independent local estimates. Further, the state-variance vector assesses the variations in these independent local state vector estimates. The elevated values in the state-variance vector indicate faults, which are identified through a comparison with a predefined threshold, i.e., $\xi_k^{(\ell)} > \gamma_{th}^{(\ell)}$. This detection threshold is further improved using an adaptive thresholding approach [27], that enhances the effectiveness of sensor fault detection in highly dynamic conditions [36], [37].

An error metric $\boldsymbol{E}_k \triangleq \left[\boldsymbol{e}_k^{(1)}, \boldsymbol{e}_k^{(2)}, \boldsymbol{e}_k^{(3)}, \ldots, \boldsymbol{e}_k^{(M)}\right] \in \mathbb{R}^{m \times M}$ is used for adaptive thresholding, where M is the total number of sensors. The vector $\boldsymbol{e}_k^{(\ell)} \in \mathbb{R}^{m \times 1}$ consists of the local state-variance entries $\xi_k^{(\ell)}$, which satisfy the condition $\xi_k^{(\ell)} < \gamma_{th}^{(\ell)}$. The moving time window length m remains constant under the non-faulty scenario and becomes variable in the presence of faults such that $m \geq m_t$, where m_t is the lower bound on the window length. Consequently, the vector $\boldsymbol{e}_k^{(\ell)}$ can be defined as $\boldsymbol{e}_k^{(\ell)} = \left[\xi_{k-m}^{(\ell)}, \xi_{k-m+1}^{(\ell)}, \ldots, \xi_{k-1}^{(\ell)}, \xi_k^{(\ell)}\right]^T$. Moreover, the mean and variance of $\boldsymbol{e}_k^{(\ell)}$ can be obtained as $\mu_{e,k}^{(\ell)} = \frac{1}{m}\sum_{j=k-m}^{k} \xi_j^{(\ell)}$ and $\sigma_{e,k}^{(\ell)} = \frac{1}{m}\sum_{j=k-m}^{k}\left(\xi_j^{(\ell)} - \frac{1}{m}\sum_{j=k-m}^{k}\left(\xi_j^{(\ell)}\right)\right)^2$, respectively. Further, the adaptive threshold using these variables can be computed as

$$\gamma_k^{(\ell)} = r\mu_{e,k}^{(\ell)} + \lambda\sigma_{e,k}^{(\ell)}, \tag{10}$$

where the tuning factors r and λ are chosen based on the use-case.

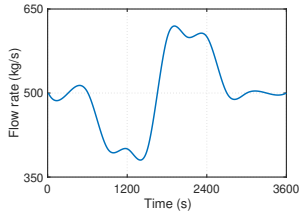
The faulty sensors/locals can be identified using the local state-variance values. If the state-variance value $\xi_k^{(\ell)}$ exceeds the threshold $\gamma_{th}^{(\ell)}$, it signifies a fault in the ℓth position sensor and the local filter processing that specific sensor measurement. Since the faulty measurements only affect their respective local state estimates, therefore, by identifying the locations of the faulty measurements one can readily identify the faulty locals using the state-variance vector [27]. The residual metric is utilized for fault isolation and to accommodate the faulty local estimates. This metric can be expressed as

$$r_{i,k}^{(\ell)} = \left|\hat{x}_{i,k|k}^{(\ell)} - \frac{1}{N}\sum_{i=1}^{N}\left(\hat{x}_{i,k|k}^{(\ell)}\right)\right|. \tag{11}$$

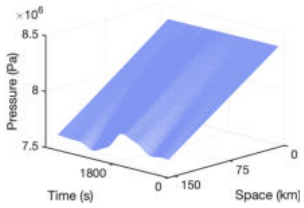
The residual entry is non-zero only when its associated local filter measurement is faulty. Otherwise, residual entries remain zero when the locals generate accurate estimates based on their respective non-faulty measurements. Hence, one can identify and eliminate the faulty estimates during the data fusion by analyzing the residual vector entries. After identifying the faulty local, its faulty local state estimate $\hat{\boldsymbol{x}}_{i,k|k}$ can be corrected using the local state residual vector $\boldsymbol{r}_{i,k}$ as

$$\hat{x}_{i,k|k}^{(\ell)} = \begin{cases} \hat{x}_{i,k|k}^{(\ell)}, & r_{i,k}^{(\ell)} < \gamma_{th}^{(\ell)} \\ \hat{x}_{i,k|k-1}^{(\ell)}, & r_{i,k}^{(\ell)} > \gamma_{th}^{(\ell)} \end{cases}, \tag{12}$$

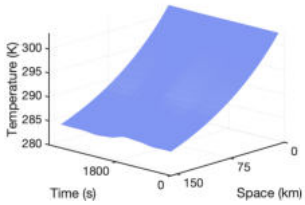
where $\gamma_{th}^{(\ell)}$ denotes the threshold to identify the faulty local state estimate. Upon identifying the fault, the ℓth entry of the



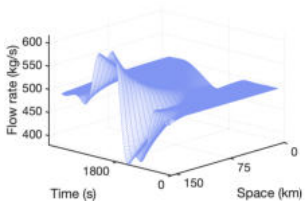
(a) Boundary condition



(b) Pressure



(c) Temperature



(d) Flow Rate

Fig. 2: Simulated data using system model.

faulty ith local filter state estimate $\hat{x}^{(\ell)}_{i,k|k}$ is substituted with the ℓth entry of the *a priori* ith local filter state estimate $\hat{x}^{(\ell)}_{i,k|k-1}$, which is not affected by the faulty local measurement. After correcting the faulty estimate, the local state estimate is combined using the information fusion.

Step 4: Information fusion in the main filter:

Using the *a posteriori* state estimates and covariance matrices of N local filters, the global *a posteriori* state estimate $\hat{\boldsymbol{x}}_{k|k}$ and the global state covariance matrix $\boldsymbol{P}_{k|k}$ corresponding to the main filter can be evaluated as [23]

$$\boldsymbol{P}_{k|k} = \left(\sum_{i=1}^{N} \boldsymbol{P}^{-1}_{i,k|k} \right)^{-1}, \hat{\boldsymbol{x}}_{k|k} = \boldsymbol{P}_{k|k} \sum_{i=1}^{N} \boldsymbol{P}^{-1}_{i,k|k} \hat{\boldsymbol{x}}_{i,k|k}. \quad (13)$$

The final state estimate $\hat{\boldsymbol{x}}_{k|k}$ is subsequently used as *a prior* information during the next iteration.

IV. Simulation Results and Discussion

We assess the performance of our proposed design using the simulated data generated by the transient flow model detailed in Section II. The parameters for the simulation are $L = 150$ km, $d = 1.4$ m, $\epsilon = 0.016$ mm, $T_s = 5^o$C, and $U = 2.84$ Wm^{-2}K^{-1}. The boundary conditions are set as $p(0,t) = 8.4$ MPa and $T(0,t) = 303.15$ K, $\dot{m}(L,t) = f(t)$, covering a time interval of $t_f \in [0, 3600$ s$]$. The spatial and temporal step sizes are $\Delta s = 7500$ m and $\Delta t = 10$ s, respectively. The simulated data without any observation noise is illustrated in Fig. 2.

To assess the state estimation of the proposed architecture, we introduce the measurement noise under three signal-to-noise ratios (SNRs): low, moderate, and high. The distribution of the measurement noise for these scenarios is presented in Table I. We consider $M = 63$ sensors and $N = 3$ locals for our experiments, where each local filter receives 21 sensor measurements. An ensemble size of $N_e = 120$ is selected for optimal performance of the proposed EnKF framework. The standard deviation of the process noise is set to be 10% lower than that of the measurement noise. Both covariance matrices $\boldsymbol{Q}_k$ and $\boldsymbol{R}_{i,k}$ are considered diagonal, where the diagonal entries represent the variances of the process and measurement noise, respectively.

Measurement noise	Pressure	Flow rate	Temperature
High SNR	$\mathcal{N}(0, 0.00005^2)$	$\mathcal{N}(0, 0.25^2)$	$\mathcal{N}(0, 0.15^2)$
Moderate SNR	$\mathcal{N}(0, 0.0005^2)$	$\mathcal{N}(0, 2.5^2)$	$\mathcal{N}(0, 1.5^2)$
Low SNR	$\mathcal{N}(0, 0.005^2)$	$\mathcal{N}(0, 10^2)$	$\mathcal{N}(0, 6^2)$

TABLE I: Distribution of measurement noise for various SNR scenarios.

SNR	Filter	Pressure (10^{-3} MPa)	Flow rate (kg/s)	Temperature (K)
High	Proposed (Partial-Distributed)	0.5701	0.8916	0.2024
	Proposed (Fully-Distributed)	0.5549	0.8722	0.1721
	Fusing UKF	0.398	0.8561	0.3220
	MM-SFDIA	0.548	2.34	2.6338
Moderate	Proposed (Partial-Distributed)	0.5299	0.8445	0.1893
	Proposed (Fully-Distributed)	0.5617	0.8802	0.1987
	Fusing UKF	0.1230	0.6463	0.3472
	MM-SFDIA	0.535	2.5235	2.9985
Low	Proposed (Partial-Distributed)	0.7926	1.9119	0.7665
	Proposed (Fully-Distributed)	1.044	1.72	0.8276
	Fusing UKF	0.227	0.8972	0.3761
	MM-SFDIA	7.8150	46	40.8613

TABLE II: Comparing the RMSE of different filters in non-faulty scenarios.

The estimation performance is evaluated in terms of the root mean square error (RMSE) in both spatial and temporal dimensions [27]. The estimation performance of the proposed technique is compared against different baselines, including the proposed fully-distributed architecture, UKF [23], and model-based multi-sensor fault detection, isolation, and accommodation (MM-SFDIA) [27], all utilizing the fully distributed filtering. In the proposed fully distributed filtering architecture, all local filters perform nonlinear time updates and linear measurement updates, while the main filter focuses on information fusion. It can be observed from the RMSE results, given in Table II, that our proposed design achieves a comparable level of estimation accuracy as that of the fully distributed technique and other baselines. Our approach, leveraging the partial-distributed state estimation, maintains comparable estimation accuracy while significantly reducing the computational requirements.

To evaluate the proposed estimation method with the SFDIA mechanism, synthetically generated faults (bias and drift [9]) are introduced in the simulated data obtained from the transient-flow model. For the adaptive thresholding, m, m_t are considered as $m = 10$ and $m_t = 5$, respectively. For performance evaluation in a multiple sensor fault scenario, three simultaneous bias and drift faults are introduced to the pressure, flow rate, and temperature sensors (indexed at $\ell = 12, 35$, and 51, respectively) for 300 s $\leq t \leq$ 500 s. As depicted in Fig. 3, our proposed architecture demonstrates superior performance compared to other techniques. The detection performance is assessed through receiver operating characteristic (ROC) plots, indicating probabilities of detection and false alarms per sample. These plots are generated by varying the tuning parameters (r and λ) of the adaptive threshold. Results

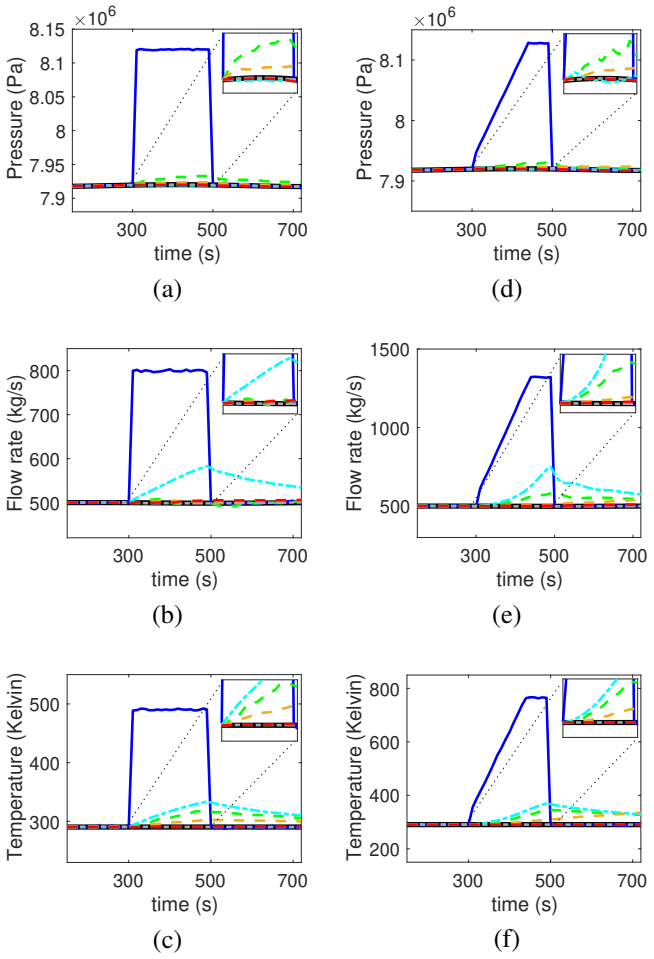


Fig. 3: State estimation techniques comparison (classic EnKF (cyan), fusing UKF (green) [23], MM-SFDIA (gray) [27], proposed without SFDIA (orange) and proposed with SFDIA (red) in the presence of three simultaneous bias and drift faults occurring at indices $\ell = 12, 35, 51$. Bias and drift fault results are depicted in plots (a)-(c) and plots (d)-(f), respectively. Actual values are indicated in black, while faulty values are shown in blue.

in Fig. 4 illustrate that our technique achieves significantly higher detection probability and considerably lower false alarm probability, especially in weak fault scenarios, compared to a recently introduced multi-sensor fault diagnosis approach MM-SFDIA [27].

Additionally, the computational complexity in terms of the execution time and the number of local filters required to address both single and multiple faults, is expressed in Table III. These results distinctly illustrate that our proposed architecture notably reduces the execution time and utilizes fewer local filters in scenarios involving both single and multiple faults. Furthermore, Table IV presents a comparison for single fault detection. The results demonstrate that the proposed approach outperforms the fusing UKF technique [23], specifically tailored for the single fault scenario.

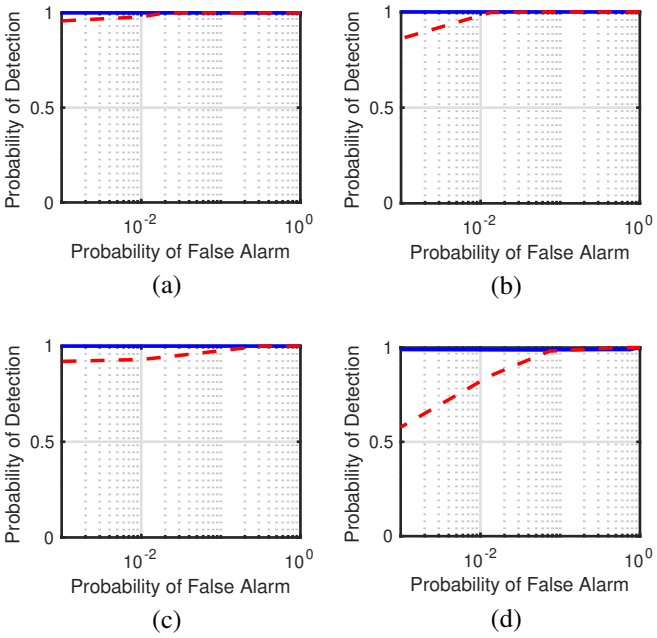


Fig. 4: ROC plots indicating the detection performance of proposed partial-distributed (blue) and MM-SFDIA [27] (red) for three simultaneous faults of different types (a) strong bias, (b) weak bias, (c) strong drift, and (d) weak drift.

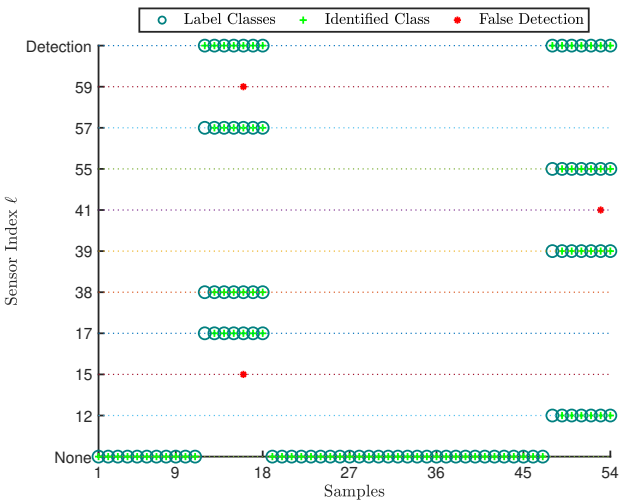


Fig. 5: Fault isolation and detection visualization of the proposed architecture.

Moreover, the fault isolation capability of the proposed method in the presence of multiple faults, with a focus on weak bias faults, is illustrated in Fig. 5. We consider 50 samples to generate the results. In the figure, ○ symbol denotes the actual fault and None signifies a scenario without any faults within the system. The falsely detected sensor fault is specified with the ∗ symbol, and the + symbol indicates the correctly identified fault. The proposed architecture effectively detects all faults with only a few instances of false detection.

Method	Sensors	1 Fault Case		Multiple 3 Faults	
		Execution Time (seconds)	Number of locals	Execution Time (seconds)	Number of locals
Proposed (Partial-Distributed)	63	**0.57s**	**3**	**0.57s**	**3**
Proposed (Fully-Distributed)	63	1.049s	**3**	1.049s	**3**
MM-SFDIA	63	3.27s	**3**	3.27s	**3**
Fusing UKF	63	27.11s	63	17955.57	41727

TABLE III: Comparing the computational complexity of the proposed partial-distributed design with the proposed fully-distributed, MM-SFDIA, and fusing UKF in terms of the execution time and the number of locals required for handling single and multiple fault cases.

Fault	Type	Proposed (Partial-Distributed)	MM-SFDIA	Fusing UKF
Pressure	Weak Bias	100	100	100
	Weak Drift	100	100	100
Flow rate	Weak Bias	100	100	100
	Weak Drift	100	85.56	83.89
Temperature	Weak Bias	100	100	97.92
	Weak Drift	100	99.72	90.83

TABLE IV: Comparing the detection accuracy (%) of the proposed partial-distributed design with MM-SFDIA, and fusing UKF for single fault detection case while considering different fault types.

V. Conclusions

This paper introduces a novel approach for state estimation and sensor fault diagnosis in natural gas pipelines experiencing transient flow. The proposed method utilizes a partial-distributed EnKF framework, separating the nonlinear computations from the local filters to the main filter, thereby significantly reducing the computational overhead. A new fault detection and isolation technique, leveraging the local state variances and local state residuals, is developed for simultaneously occuring multiple sensor faults. Further, adaptive thresholding is introduced to enhance the fault detection accuracy. Simulations demonstrate the effectiveness of the proposed architecture in accurately detecting and isolating multiple simultaneous sensor faults. Future work aims to address process-related faults, fault detection in pipelines with varying operating conditions, and design of control laws/observers for governing PDEs.

References

[1] L. Hou, D. Wang, B. Du, X. Qian, and M. Yuan, "Gas concentration detection via multi-channeled air sampling method," *Sensor Review*, vol. 37, no. 2, pp. 187–195, 2017.

[2] Q. Tan, G. Feng, H. Yuan, G. Su, M. Fu, and Y. Zhu, "Applied research for the gas pipeline installation and the monitoring method for the safety of the adjacent under-ground spaces," *Journal of safety and environment*, vol. 19, pp. 902–908, 2019.

[3] T. B. Quy and J.-M. Kim, "Real-time leak detection for a gas pipeline using ak-nn classifier and hybrid ae features," *Sensors*, vol. 21, no. 2, p. 367, 2021.

[4] J. Wan, Y. Yu, Y. Wu, R. Feng, and N. Yu, "Hierarchical leak detection and localization method in natural gas pipeline monitoring sensor networks," *Sensors*, vol. 12, no. 1, pp. 189–214, 2011.

[5] R. Xiao, Q. Hu, and J. Li, "Leak detection of gas pipelines using acoustic signals based on wavelet transform and support vector machine," *Measurement*, vol. 146, pp. 479–489, 2019.

[6] A. Chawla, Y. Arellano, M. V. Johansson, H. Darvishi, K. Shaneen, M. Vitali, F. Finotti, and P. S. Rossi, "Iot-based monitoring in carbon capture and storage systems," *IEEE Internet of Things Magazine*, vol. 5, no. 4, pp. 106–111, 2022.

[7] Y. Feng and H.-X. Li, "Detection and spatial identification of fault for parabolic distributed parameter systems," *IEEE Transaction on Industrial Electronics*, vol. 66, no. 9, pp. 7300–7309, 2018.

[8] J. Yang, F. Zhu, X. Wang, and X. Bu, "Robust sliding-mode observer-based sensor fault estimation, actuator fault detection and isolation for uncertain nonlinear systems," *International Journal of Control, Automation and Systems*, vol. 13, no. 5, pp. 1037–1046, 2015.

[9] H. Darvishi, D. Ciuonzo, E. R. Eide, and P. Salvo Rossi, "Sensor-fault detection, isolation and accommodation for digital twins via modular data-driven architecture," *IEEE Sensors Journal*, vol. 21, no. 4, pp. 4827–4838, 2021.

[10] H. Darvishi, D. Ciuonzo, and P. Salvo Rossi, "A machine-learning architecture for sensor fault detection, isolation, and accommodation in digital twins," *IEEE Sensors Journal*, vol. 23, no. 3, pp. 2522–2538, 2023.

[11] F. Serdio, E. Lughofer, K. Pichler, T. Buchegger, and H. Efendic, "Residual-based fault detection using soft computing techniques for condition monitoring at rolling mills," *Information Sciences*, vol. 259, pp. 304–320, 2014.

[12] X. Wei, M. Verhaegen, and T. van Engelen, "Sensor fault detection and isolation for wind turbines based on subspace identification and kalman filter techniques," *International Journal of Adaptive Control and Signal Processing*, vol. 24, no. 8, pp. 687–707, 2010.

[13] W. El Sayed, M. Abd El Geliel, and A. Lotfy, "Fault diagnosis of PMSG stator inter-turn fault using extended Kalman filter and unscented Kalman filter," *Energies*, vol. 13, no. 11, p. 2972, 2020.

[14] L. Yan, H. Zhang, X. Dong, Q. Zhou, H. Chen, and C. Tan, "Unscented Kalman-filter-based simultaneous diagnostic scheme for gas-turbine gas path and sensor faults," *Measurement Science and Technology*, vol. 32, no. 9, p. 095905, 2021.

[15] D. Mori, H. Sugiura, and Y. Hattori, "Adaptive sensor fault detection and isolation using unscented kalman filter for vehicle positioning," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*. IEEE, 2019, pp. 1298–1304.

[16] B. Du, Z. Shi, J. Song, H. Wang, and L. Han, "A fault-tolerant data fusion method of MEMS Redundant Gyro system based on weighted distributed Kalman filtering," *Micromachines*, vol. 10, no. 5, p. 278, 2019.

[17] Y. Luo, Y. Liu, W. Yang, J. Zhou, and T. Lv, "Distributed filtering algorithm based on local outlier factor under data integrity attacks," *Journal of the Franklin Institute*, 2023.

[18] N. Sadeghzadeh Nokhodberiz and J. Poshtan, "Belief consensus–based distributed particle filters for fault diagnosis of non-linear distributed systems," *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, vol. 228, no. 3, pp. 123–137, 2014.

[19] P. Jin, X. Zhou, C. Wang, J. Huang, W. Zhou, and F. Lu, "A novel distributed Kalman filtering for health state recognition of aero-engine components in networked control systems," *Nonlinear Dynamics*, pp. 1–19, 2022.

[20] K. Salahshoor, M. Mosallaei, and M. Bayat, "Centralized and decen-

tralized process and sensor fault monitoring using data fusion based on adaptive extended Kalman filter algorithm," *Measurement*, vol. 41, no. 10, pp. 1059–1076, 2008.
[21] K. Geng and N. Chulin, "Applications of multi-height sensors data fusion and fault-tolerant Kalman filter in integrated navigation system of UAV," *Procedia Computer Science*, vol. 103, pp. 231–238, 2017.
[22] F. Lu, Y. Wang, J. Huang, Y. Huang, and X. Qiu, "Fusing unscented Kalman filter for performance monitoring and fault accommodation in gas turbine," *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, vol. 232, no. 3, pp. 556–570, 2018.
[23] A. Vafamand, B. Moshiri, and N. Vafamand, "Fusing unscented Kalman filter to detect and isolate sensor faults in DC microgrids with CPLs," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–8, 2021.
[24] K. Shaheen, A. Chawla, F. E. Uilhoorn, and P. S. Rossi, "Model-based sensor-fault detection and isolation in natural-gas pipelines for transient flow," in *2023 IEEE SENSORS*, 2023, pp. 1–4.
[25] k. shaheen, A. Chawla, F. Uilhoorn, and P. Salvo Rossi, "Sensor-fault detection, isolation and accommodation for natural-gas pipelines under transient flow," *Under review*, 2023.
[26] F. Lu, Z. Li, J. Huang, and M. Jia, "Hybrid state estimation for aircraft engine anomaly detection and fault accommodation," *AIAA Journal*, vol. 58, no. 4, pp. 1748–1762, 2020.
[27] K. Shaheen, A. Chawla, F. E. Uilhoorn, and P. S. Rossi, "Model-based architecture for multi-sensor fault detection, isolation and accommodation in natural-gas pipelines," *IEEE Sensors Journal*, pp. 1–1, 2023.
[28] H. Fang, N. Tian, Y. Wang, M. Zhou, and M. A. Haile, "Nonlinear bayesian estimation: From kalman filtering to a broader horizon," *IEEE/CAA Journal of Automatica Sinica*, vol. 5, no. 2, pp. 401–417, 2018.
[29] F. Uilhoorn, "Comparison of Bayesian estimation methods for modeling flow transients in gas pipelines," *Journal of Natural Gas Science and Engineering*, vol. 38, pp. 159–170, 2017.
[30] F. E. Uilhoorn, "A particle filter-based framework for real-time state estimation of a non-linear hyperbolic PDE system describing transient flows in CO2 pipelines," *Computers & Mathematics with Applications*, vol. 68, no. 12, pp. 1991–2004, 2014.
[31] F. Uilhoorn, "Dynamic behaviour of non-isothermal compressible natural gases mixed with hydrogen in pipelines," *International journal of hydrogen energy*, vol. 34, no. 16, pp. 6722–6729, 2009.
[32] C. F. Colebrook, T. Blench, H. Chatley, E. Essex, J. Finniecome, G. Lacey, J. Williamson, and G. Macdonald, "Correspondence. turbulent flow in pipes, with particular reference to the transition region between the smooth and rough pipe laws.(includes plates)." *Journal of the Institution of Civil engineers*, vol. 12, no. 8, pp. 393–422, 1939.
[33] E. W. Lemmon and R. T. Jacobsen, "Viscosity and thermal conductivity equations for nitrogen, oxygen, argon, and air," *International Journal of Thermophysics*, vol. 25, pp. 21–69, 2004.
[34] W. E. Schiesser, *The numerical method of lines: Integration of partial differential equations*. Elsevier, 2012.
[35] R. Courant, "Uber die partiellen Differenzengleichungen der mathematischen Physik," *Math Ann*, vol. 100, no. 1, pp. 32–74, 1928.
[36] K. Patan, *Artificial neural networks for the modelling and fault diagnosis of technical processes*. Springer, 2008.
[37] S. Amirkhani, A. Chaibakhsh, and A. Ghaffari, "Nonlinear robust fault diagnosis of power plant gas turbine using Monte Carlo-based adaptive threshold approach," *ISA transactions*, vol. 100, pp. 171–184, 2020.