Partial-Distributed Architecture for Multisensor Fault Detection, Isolation, and Accommodation in Hydrogen-Blended Natural Gas Pipelines

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

***Abstract*—This article investigates an innovative state estimation technique implemented within an advanced distributed framework, aimed at reducing computational complexity while detecting multiple sensor faults in hydrogen-blended natural gas pipelines. The novel distributed estimation technique is based on the ensemble Kalman filter (EnKF) and is referred to as partial-distributed multisensor fault detection, isolation, and accommodation. The architecture includes a set of local EnKFs and an information fusion center. These local filters operate simultaneously to generate unique local state estimates based on a distinct set of sensor measurements, which are subsequently transmitted to the information fusion center for the computation of fault-free state estimates. To reduce computational complexity, the partially distributed approach segregates nonlinear computations from the local filters and delegates them to the main filter. Additionally, a fault diagnosis strategy is developed based on local state residuals. Since each local filter generates a distinct local state estimate based on its unique set of sensor measurements, comparing the local state residual against a threshold facilitates the identification and isolation of faulty sensors. Furthermore, an adaptive thresholding approach is incorporated to facilitate effective fault identification and isolation. The proposed technique has proven to be effective in highly nonlinear, and high-dimensional systems with simultaneous multiple sensor faults. The effectiveness of the proposed approach is demonstrated through extensive simulations and comparative analyses.**

***Index Terms*—Data fusion, ensemble Kalman filter (EnKF), fault diagnosis, hydrogen-blended natural gas pipelines, multiple sensor faults, partial-distributed architecture.**

I. Introduction

THE CONTINUOUS surveillance of gas pipelines is of paramount importance to guarantee the overall safety and reliability of the entire system. With the progression of digital technologies, several monitoring systems for urban gas pipelines have emerged for potential leak identification [1], [2]. These systems typically employ a variety of gas monitoring sensors to continuously observe the gas pipelines, aiming to identify potential leaks and enhance the reliability of the pipeline infrastructures [3], [4], [5]. However, it is crucial to acknowledge that the sensors installed in these pipelines can occasionally become faulty due to factors, such as harsh environmental conditions, aging, improper calibration, and hardware failures, or even communication errors. The sensor faults causing incorrect measurements could result in delayed or overlooked leak detection, potentially leading to wrong decisions with potentially severe consequences [6]. Hence, there is a critical need to develop effective techniques that can identify sensor faults within gas-monitoring systems in a timely manner. Moreover, the growing adoption of hydrogen-blended natural gas for environmental-friendly heating and power generation introduces a new challenge for gas monitoring sensors, necessitating special attention and consideration that have not been previously addressed.

The sensor fault diagnosis techniques can generally be grouped into two categories: 1) data-driven [7] and 2) model-based [8] approaches. Data-driven methods utilize real-time/historical data to diagnose sensor faults [9], [10], [11], [12]. However, their reliance on extensive recorded data streams necessitates a sizeable database, and their substantial computational requirements render them impractical for online applications. On the contrary, model-based methods rely on the assumption that an accurate system model and its parameters are available. They entail comparing actual process measurements with those predicted by the model, generating a residual signal for detecting faults [13], [14], [15]. These methods outperform the data-driven approaches when an accurate mathematical system model exists, offering better real-time performance.

The current state-of-the-art approaches have largely overlooked the substantial challenge of managing multiple sensor failures in large-scale and highly nonlinear systems. Moreover, the existing distributed estimation methods pose a significant challenge due to their high computational complexity, especially with large-scale systems comprising numerous subsystems or local filters. In such systems, each subsystem typically involves repetitive nonlinear calculations related to the system model, which significantly increases the overall

Manuscript received 2 July 2024; accepted 24 July 2024. Date of publication 29 July 2024; date of current version 24 October 2024. This work was supported in part by the Research Council of Norway under the Project SIGNIFY within the IKTPLUSS Framework under Project 311902. *(Corresponding author: Khadija Shaheen.)*

Khadija Shaheen and Apoorva 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).

Ferdinand Evert Uilhoorn is with the Department of Gas Engineering, Warsaw University of Technology, 00-661 Warszawa, Poland (e-mail: ferdinand.uilhoorn@pw.edu.pl).

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

Digital Object Identifier 10.1109/JIOT.2024.3435413

computational load. To tackle these challenges, it is crucial to develop new distributed architectures that minimize redundant nonlinear computations and possess the capability to efficiently handle multiple sensor faults while ensuring a low computational complexity.

A. Related Works

Several techniques are available within the model-based approaches, such as parity space [16], parameter estimation [17], and state estimation [18], [19]. Among these model-based methods, the Kalman filter (KF) has been extensively preferred due to its robustness and effectiveness, providing optimal state estimation, particularly for linear systems affected by Gaussian noise [20]. Its optimality in handling linear systems has spurred the development of various modifications to handle state estimation in nonlinear systems, including the extended KF (EKF) [21], the unscented KF (UKF) [22], the ensemble KF (EnKF) [23], and the cubature KF (CKF) [24].

In recent decades, decentralized KF-based architectures have gained significant importance to face limitations experienced by centralized KF-based architectures, such as hardware constraints, communication interference, time-varying characteristics, and other design factors [22], [25], [26]. More specifically, distributed architectures alleviate these challenges by dividing the computation among several multiple local filters: an approach enhancing decision-making while reducing the overall computational load [27], [28], [29], [30]. Distributed architectures relying on information fusion are often preferred for effectively merging outputs from numerous local filters within large-scale systems. These techniques ensure high accuracy and enhance robust fault detection while minimizing communication load [31], [32].

Centralized and distributed architectures based on EKFs are developed and compared in [33], with the former exhibiting high estimation accuracy and poor robustness to sensor faults, while the latter struggling with nonlinearities and multiple sensor faults. Robustness with respect to measurement delays and losses was handled by the distributed federated KF-based fusion method discussed in [34], which requires however significant computational power and memory in large-scale systems and still cannot deal with highly nonlinear systems.

Data fusion techniques based on redundancy for detecting multiple faults (both hardware and software) are explored in [35], [36], and [37], but their complexity significantly increases with the number of multiple faults to be addressed. In [31], a sensor-fusion approach utilizing multiple KFs is presented for nonlinear systems, where each filter is tailored for a specific defect, however, this method has significant computational cost limitations. Wasserstein average-consensus classification was introduced in [38] to deal with multiple-sensor faults, but its use is limited to linear systems. Furthermore, a trust-based distributed KF architecture is proposed in [39] to enhance resilience against cyber attacks. Doostmohammadian and Meskin [40] proposed a decentralized technique, where local estimators perform preliminary fault detection and share information with the neighboring nodes, however, this method does not account for nonlinear systems. Moreover, the consensus-based methods, where the nodes are interconnected and share estimates, require robust communication links among local filters. This poses significant challenges for large-scale systems and is beyond the scope of our work.

An EKF-based partial distributed architecture is proposed in [42] to identify sensor and process faults within a gas turbine engine. Also, our previous works [43], [44] developed a model-based sensor fault detection, isolation and accommodation (SFDIA) framework for natural gas pipelines still focusing on single fault detection. Differently, a sensor-fusion technique based on UKF is explored in [45] for monitoring a gas turbine engine, considering four different combinations of local filters for the sensors installed on the gas turbine. Although focusing on typical nonlinearities present in process engineering, all these methods are primarily designed for single-fault scenarios and becomes computationally intensive and poor performing when handling multiple sensor faults. Finally, our recent work [40] explored the detection of simultaneous multiple sensor faults using an architecture based on distributed EnKFs, however, the architecture encounters substantial computational complexity owing to a fully distributed structure.

B. Contributions of This Article

It is apparent that methods effectively handling multiple-sensor faults in large-scale nonlinear systems, as often required in process engineering and industrial applications, are not available. Inspired by the works in [45] and [40], the architecture proposed here introduces a novel partially distributed (PD) filtering framework handling multiple-sensor faults within a highly nonlinear, complex, and large-scale distributed system of natural gas pipeline while significantly reducing the computations by separating the nonlinear calculations from the local filters and relocating them to the main filter.

More specifically, we propose a novel partial-distributed EnKF-based architecture designed to alleviate the computational challenges associated with nonlinear state estimation and handle faults of multiple sensors in natural gas pipelines. The proposed framework leverages the partial-distributed filtering structure that segregates the nonlinear computations from the local filters and transfers them to the main filter. The main filter is responsible for managing the time update and information fusion, while the local filters focus on the measurement updates. This approach significantly reduces the computational complexity of the overall system while offering distinct advantages in high-dimensionality and large-scale systems via the EnKF-based architecture. Further, a novel fault-tolerant approach based on local state variance and residual is introduced. This approach involves identifying high values in these metrics during faulty conditions, aiding in the isolation of faulty state estimates from the information fusion scheme. Also, an adaptive thresholding technique dynamically adjusts according to the evolving system conditions, thereby improving the accuracy of the fault detection. The final global estimate is obtained using the information fusion that merges the nonfaulty outputs of all the local filters.

TABLE I
COMPARING THE PROPOSED WORK WITH THE EXISTING STATE-OF-THE-ART WORKS

Works	Multiple faults	Distributed /Partial	Filter type	Computation cost	Adaptive Thresholding	Application	Large scale-system	Random weak faults
[33]	✗	Distributed	EKF	↑ High	✗	Continuous stirred tank reactor	✗	✗
[31]	✗	Distributed	KF	↑ High	✗	Navigation system	✗	✗
[34]	✓	Distributed	KF	↑ High	✗	Target tracking	✗	✗
[39]	✓	Distributed	KF	↑ High	✗	Target tracking	✗	✗
[40]	✓	Distributed	KF	↑ High	✓	Communication Network	✗	✗
[44]	✗	Distributed	EKF, UKF	↑ High	✗	Gas turbine engine	✗	✗
[32]	✗	Distributed	UKF	↑ High	✗	Micro grid	✗	✗
[42]	✗	Distributed	UKF	↑ High	✗	Natural gas pipeline	✓	✓
[43]	✗	Distributed	EKF,EnKF,UKF	↑ High	✗	Natural gas pipeline	✓	✓
[45]	✓	Distributed	EnKF	↑ High	✓	Natural gas pipeline	✓	✓
[41]	✗	Partial	EKF	↓ Low	✗	Gas turbine	✗	✗
Proposed	✓	Partial	EnKF	↓ Low	✓	Hydrogen-blended natural gas pipeline	✓	✓

Moreover, we consider the transportation of hydrogen-blended natural gas through pipelines, a practice recently endorsed by natural gas utilities [46]. The thermodynamic properties and flow conditions in pipelines carrying hydrogen-blended natural gas differ from those handling only natural gas. Additionally, we employ more sensors and highly nonlinear flow conditions that increase the complexity of the system dynamics in comparison to [43] and [45]. Under these complex conditions, our proposed technique offers superior estimation and fault detection performance, coupled with low execution time compared to the existing methodologies.

Table I presents a comparative qualitative analysis of our proposed approach against the relevant state-of-the-art fault diagnosis techniques. The key contributions of this article are as follows.

1) We explore for the first time an innovative model-based sensor fault detection tailored for hydrogen-blended natural gas pipelines undergoing transient flow. Our study delves into a comprehensive flow model for hydrogen-blended natural gas, encompassing its numerical solution and emphasizing its practical applicability in fault diagnosis (previously utilized solely for state estimation [46]).
2) A novel architecture based on a PD EnKF-based framework is proposed for nonlinear state estimation. The architecture effectively separates the nonlinear computations from the local filters and redirects them to the main filter while achieving exceptionally low computational cost.
3) Further, a unique fault detection and isolation framework based on local state residual is introduced to effectively handle simultaneous multiple-sensor faults in large-scale distributed systems. Additionally, an adaptive thresholding technique is introduced to improve the fault detection accuracy.
4) The performance of the proposed architecture is assessed under the influence of synthetically induced weak bias and drift sensor faults. These faults simulate abrupt (hard) faults and gradually emerging (soft) faults.

The remainder of this article is organized as follows. Section II introduces the transient flow model. Section III elaborates on the proposed architecture and the adaptive thresholding technique. Simulation results are presented and analyzed in Section IV to substantiate the effectiveness of the proposed architecture. Section V concludes this article with some final remarks and possible future research directions.

II. TRANSIENT FLOW MODEL FOR HYDROGEN-BLENDED NATURAL GAS PIPELINE

The transient flow of hydrogen-blended natural gas in pipelines can be represented by a set of hyperbolic partial differential equations (PDEs), given as [47]

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

where t and s denote the time and spatial domain, respectively. The spatio–temporal behavior can be collectively described as $\Omega = \{(s,t) : 0 \le s \le L, 0 \le t \le t_f\}$, where L is the length of the pipeline and t_f indicates the time duration. Further, the state vector $\boldsymbol{x} = [p, \dot{m}, T]^{\mathrm{T}}$ includes pressure (p), flow rate ($\dot{m}$), and temperature (T), while $(\cdot)^{\mathrm{T}}$ is the transpose operator.

The matrix $\boldsymbol{A}(\boldsymbol{x}) \in \mathbb{R}^{3\times 3}$, provided in (2), shown at the bottom of the page, represents the coefficient matrix [48], [49], and the vector $\boldsymbol{\zeta}(\boldsymbol{x}) \in \mathbb{R}^{3\times 1}$ is defined as

$$\boldsymbol{\zeta}(\boldsymbol{x}) = \left[-\frac{a_s^2 a_1 (Aqp + RT\dot{m}wz)}{A^2 TC_p p} \quad w \quad -\frac{a_s^2 a_2 (Aqp + RT\dot{m}wz)}{A^2 C_p p^2} \right]^{\mathrm{T}}$$

where $a_1 = 1 + [T/z](\partial z/\partial T)_p$, $a_2 = 1 - [p/z](\partial z/\partial p)_T$. The symbols a_s z, R, and C_p represent the isentropic wave speed, the gas compressibility factor, the ideal gas constant, and the specific heat at constant pressure, respectively [47], [50], [51].

$$\boldsymbol{A}(\boldsymbol{x}) = \begin{bmatrix} -\frac{\dot{m}(a_s^2 a_2 - RTz)}{Ap} & \frac{a_s^2}{A} & \frac{a_s^2 \dot{m} a_1}{AT} \\ A - \frac{a_s^2 a_2^2 C_p \dot{m}^2 - Ra_s^2 a_1^2 a_2 \dot{m}^2 z}{AC_p p^2} & \frac{\dot{m}(a_2 C_p a_s^2 - Rza_s^2 a_1^2 + RTC_p z)}{AC_p p} & \frac{a_s^2 a_1 \dot{m}^2 (a_2 C_p - Ra_1^2 z)}{ATC_p p} \\ -\frac{RTa_s^2 a_1 a_2 \dot{m} z}{AC_p p^2} & \frac{RTa_s^2 a_1 z}{AC_p p} & \frac{R\dot{m}z(a_s^2 a_1^2 + TC_p)}{AC_p p} \end{bmatrix} \tag{2}$$

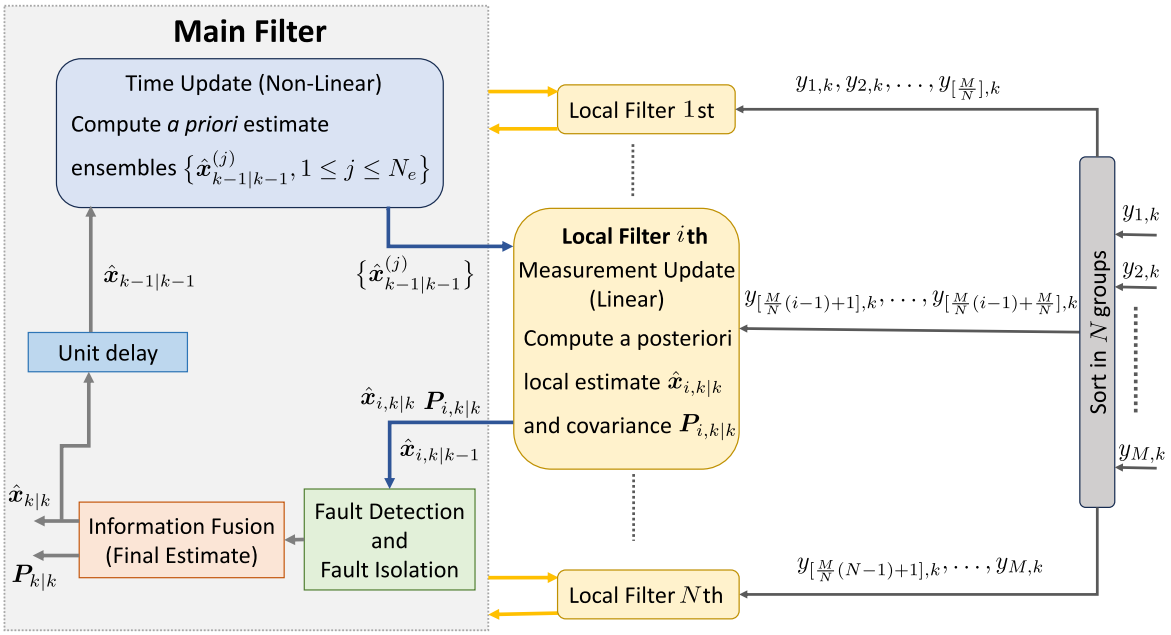

Fig. 1. Block diagram of the proposed partial-distributed architecture.

Exploiting the numerical method of lines [52] which relies on spatial discretization, a 5-point, fourth-order finite difference method is employed to convert the system of PDEs in (1) into a set of ordinary differential equations (ODEs). This approximation introduces an error of order $\mathcal{O}(\Delta s^4)$, with Δs being the spatial step size. The modified state vector $\boldsymbol{x}(t) \in \mathbb{R}^{3n\times 1}$ can be expressed as

$$\boldsymbol{x}(t) = \left[p_1(t), \ldots, p_n(t), \dot{m}_1(t), \ldots, \dot{m}_n(t), T_1(t), \ldots, T_n(t)\right]^{\mathrm{T}} \tag{3}$$

with $p_i(t)$, $\dot{m}_i(t)$, and $T_i(t)$ denoting pressure, flow rate, and temperature to the ith spatial node. Moreover, the system of ODEs can be formulated as

$$\frac{d\boldsymbol{x}(t)}{dt} = \boldsymbol{B}(\boldsymbol{x})\boldsymbol{D}\boldsymbol{x}(t) - \boldsymbol{c}(\boldsymbol{x}, t) \triangleq \boldsymbol{\varphi}(\boldsymbol{x}(t), t) \tag{4}$$

where $\boldsymbol{B}(\boldsymbol{x}) \in \mathbb{R}^{3n\times 3n}$ is the assembled matrix of $\boldsymbol{A}(\boldsymbol{x})$, is a full rank matrix and completely observable $\boldsymbol{c}(\boldsymbol{x}, t) \in \mathbb{R}^{3n\times 1}$ is the assembled column vector of $\boldsymbol{\zeta}(\boldsymbol{x})$, and

$$\boldsymbol{D} = -\frac{1}{12\Delta s}\begin{bmatrix} -25 & 48 & -36 & 16 & -3 & \cdots & 0 \\ -3 & -10 & 18 & -6 & 1 & \cdots & 0 \\ 1 & -8 & 0 & 8 & -1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & -8 & 0 & 8 & -1 \\ 0 & \cdots & -1 & 6 & -18 & 10 & 3 \\ 0 & \cdots & 3 & -16 & 36 & -48 & 25 \end{bmatrix}.$$

The system of ODEs in (4) can be further solved using an effective numerical technique known as the fourth-order Runge–Kutta (RK4) integrator. When discretized using a constant time step (Δt), the system of ODEs can be expressed as a state-space model and its evolved solution at the time instant $t_k = k \cdot \Delta t$, denoted $\boldsymbol{x}_k = \boldsymbol{x}(t_k)$ can be described as

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

where $\boldsymbol{k}_1 = \boldsymbol{\varphi}(\boldsymbol{x}_k, k\Delta t)$, $\boldsymbol{k}_2 = \boldsymbol{\varphi}(\boldsymbol{x}_k + \boldsymbol{k}_1[\Delta t/2], (2k+1)[\Delta t/2])$, $\boldsymbol{k}_3 = \boldsymbol{\varphi}(\boldsymbol{x}_k + \boldsymbol{k}_2[\Delta t/2], (2k+1)[\Delta t/2])$, and $\boldsymbol{k}_4 = \boldsymbol{\varphi}(\boldsymbol{x}_k + \boldsymbol{k}_3\Delta t, (k+1)\Delta t)$.

To maintain numerical stability, it is important to satisfy the Courant–Friedrichs–Lewy (CFL) condition [53], expressed as

$$\frac{\Delta t}{\Delta s} \le \frac{1}{|v| + a_s}. \tag{5}$$

III. Partially-Distributed Multisensor Fault Detection, Isolation, and Accommodation Architecture

The proposed design exploits a PD filtering structure, separating the nonlinear computations from the local filters and assigning them to the main filter [42]. Here, the main filter handles the time update and the information fusion, while the local filters focus on the measurement updates. This strategy significantly enhances the computational efficiency and reduces the system complexity.

To implement the proposed architecture, all sensor measurements are initially grouped into distinct sets of measurements. Each subset of measurements is assigned to a particular local filter, as depicted in Fig. 1. Subsequently, the state estimation is performed by the proposed partial-distributed architecture using the following steps: 1) first, the nonlinear computation (time update) is performed in the main filter; 2) the linear local filters execute the measurement updates and estimate the local state vectors and their associated covariance matrices using their specific subset of measurements; 3) faulty estimates are detected by comparing the state residual vector against an adaptive threshold and, once a fault is identified, the faulty estimate is replaced by a nonfaulty *a piori* estimate; and

4) these corrected local estimates are utilized to generate the global estimates in the information mixture, which are fed to the main filter for initialization during the subsequent iteration. In the following sections, the various steps employed in implementing the proposed partial-distributed filter are discussed in detail.

A. Grouping of Sensors

Initially, the number of locals and the number of measurements allocated to each local filter are defined. For grouping, we assume that each local filter is assigned a unique/nonrepetitive set of measurements. Here, we are considering a total of M sensors and N local filters such that each local filter receives M/N[1] sensor measurements as input. Consequently, each local filter independently generates a state estimate using its specific set of measurements. Additionally, the grouping technique guarantees that in the absence of sensor faults, the estimated local state vectors from various filters are similar. Hence, it becomes feasible to examine the deviations among the independent local state vector estimates. This methodology is highly effective in identifying faulty sensors, making the proposed architecture a resilient and reliable solution for fault diagnosis. It is worth noting that this grouping approach is most efficient when the sensor measurements are uncorrelated and independent. In the following section, the design of the proposed partial-distributed EnKF-based filter is discussed.

B. Partially-Distributed Filter Design

The nonlinear state-space model can be formulated 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 $\boldsymbol{f}(\cdot,\cdot) : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \rightarrow \mathbb{R}^{n_x}$ is the nonlinear mapping associated with the main filter, $\boldsymbol{x}_k \in \mathbb{R}^{n_x \times 1}$ denotes the state vector and $\boldsymbol{v}_k \in \mathbb{R}^{n_x \times 1} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_k)$ is the process noise. Further, $\boldsymbol{y}_{i,k} \in \mathbb{R}^{n_y \times 1}$ and $\boldsymbol{n}_{i,k} \in \mathbb{R}^{n_y \times 1} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_{i,k})$ represent the measurement and the measurement noise corresponding to the ith local filter, respectively, where $i = 1, 2, \ldots, N$, while we denote with $\boldsymbol{y}_k = [y_{1,k}, \ldots, y_{M,k}]^{\mathrm{T}}$ the collection of the measurements from all the sensors at the kth time slot. The input vector $\boldsymbol{u}_{k-1} \in \mathbb{R}^{n_u \times 1}$ is defined as $\boldsymbol{u}_{k-1} = [\boldsymbol{u}_{in}^{\mathrm{T}} \boldsymbol{u}_{bc,k-1}^{\mathrm{T}}]^{\mathrm{T}}$, which includes both initial and boundary conditions. Additionally, the mapping $\boldsymbol{h}_i(\cdot,\cdot) : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \rightarrow \mathbb{R}^{n_y}$ characterizes the measurement model of the ith local filter. In our case, a linear measurement model is considered such that the mapping $\boldsymbol{h}_i(\boldsymbol{x}_k, \boldsymbol{u}_k)$ is defined as $\boldsymbol{h}_i(\boldsymbol{x}_k, \boldsymbol{u}_k) = \boldsymbol{H}_i \boldsymbol{x}_k$. Here, $\boldsymbol{H}_i \in \mathbb{R}^{n_y \times n_x}$ denotes the observation matrix of the ith local filter, as described in [54]. The primary objective is to facilitate dimensionality reduction of the state vector through one-to-one mapping, depending upon the number of measurements allocated to each local filter.

For state vector estimation, we propose a novel partial-distributed EnKF-based design. The proposed algorithm is derived from the EnKF, a prominent statistical approach for approximating probability distributions by directly drawing samples from specific distributions. To address computational constraints, the nonlinear computations are shifted from the local filters to the main filter, i.e., the measurement updates are carried out in the local filters, while the time update and information fusion are performed in the main filter. This method serves as a robust tool for state estimation in high-dimensional nonlinear systems. Next, the different stages encompassed in the state estimation process of the PD EnKF architecture are outlined.

1) Initialization: First, the state vector estimate $\hat{\boldsymbol{x}}_{0|0}$ is initialized depending on the use case.

2) Nonlinear Computation in the Main Filter: At the main filter, to approximately represent the conditional probability distribution $p(\boldsymbol{x}_{k-1}|\mathbb{Y}_{k-1})$, being $\mathbb{Y}_{k-1} = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_{k-1}\}$, an ensemble of N_e samples, denoted $\{\hat{\boldsymbol{x}}_{k-1|k-1}^{(j)}, 1 \leq j \leq N_e\}$, is generated according to

$$\hat{\boldsymbol{x}}_{k-1|k-1}^{(j)} = \hat{\boldsymbol{x}}_{k-1|k-1} + \boldsymbol{v}_{k-1}^{(j)}, \quad j = 1, 2, \ldots, N_e \tag{6}$$

where $\boldsymbol{v}_k^{(j)}$ is noise samples generated according to the process noise distribution $\mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_k)$. Using the nonlinear mapping, the *a priori* ensemble representing $p(\boldsymbol{x}_k|\mathbb{Y}_{k-1})$, denoted $\{\hat{\boldsymbol{x}}_{k|k-1}^{(j)}, 1 \leq j \leq N_e\}$, can be produced as

$$\hat{\boldsymbol{x}}_{k|k-1}^{(j)} = \boldsymbol{f}\left(\hat{\boldsymbol{x}}_{k-1|k-1}^{(j)}, \boldsymbol{u}_{k-1}^{(j)}\right) + \boldsymbol{v}_k^{(j)}. \tag{7}$$

3) Linear Computations in the Local Filters: At the ith local filter, to approximately represent $p(\boldsymbol{y}_{i,k-1}|\mathbb{Y}_{i,k-1})$, being $\mathbb{Y}_{i,k-1} = \{\boldsymbol{y}_{i,1}, \boldsymbol{y}_{i,2}, \ldots, \boldsymbol{y}_{i,k-1}\}$, an ensemble of N_e measurement samples, denoted $\{\hat{\boldsymbol{y}}_{i,k|k-1}^{(j)}, 1 \leq j \leq N_e\}$, is generated according to

$$\hat{\boldsymbol{y}}_{i,k|k-1}^{(j)} = \boldsymbol{H}_i \hat{\boldsymbol{x}}_{i,k|k-1}^{(j)} + \boldsymbol{n}_{i,k}^{(j)}, \quad j = 1, 2, \ldots, N_e \tag{8}$$

where $\boldsymbol{n}_{i,k}^{(j)}$ is noise samples generated according to the measurement noise distribution $\mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_{i,k})$.

The sample mean and covariance of the measurement ensemble $\{\hat{\boldsymbol{y}}_{i,k|k-1}^{(j)}, 1 \leq j \leq N_e\}$ can be computed as

$$\hat{\boldsymbol{y}}_{i,k|k-1} = \frac{1}{N_e} \sum_{j=1}^{N_e} \hat{\boldsymbol{y}}_{i,k|k-1}^{(j)} \tag{9}$$

$$\boldsymbol{P}_{i,k|k-1}^{y} = \frac{1}{N_e - 1} \boldsymbol{E}_{i,k|k-1}^{y} \left(\boldsymbol{E}_{i,k|k-1}^{y}\right)^{\mathrm{T}} \tag{10}$$

where $\boldsymbol{E}_{i,k|k-1}^{y} = [(\hat{\boldsymbol{y}}_{i,k|k-1}^{(1)} - \hat{\boldsymbol{y}}_{i,k|k-1}), \ldots, (\hat{\boldsymbol{y}}_{i,k|k-1}^{(N_e)} - \hat{\boldsymbol{y}}_{i,k|k-1})]$. Similarly, the sample mean and covariance of the *a priori* ensemble can be computed as

$$\hat{\boldsymbol{x}}_{i,k|k-1} = \frac{1}{N_e} \sum_{j=1}^{N_e} \hat{\boldsymbol{x}}_{i,k|k-1}^{(j)} \tag{11}$$

$$\boldsymbol{P}_{i,k|k-1}^{x} = \frac{1}{N_e - 1} \boldsymbol{E}_{i,k|k-1}^{x} \left(\boldsymbol{E}_{i,k|k-1}^{x}\right)^{\mathrm{T}} \tag{12}$$

where $\boldsymbol{E}_{i,k|k-1}^{x} = [(\hat{\boldsymbol{x}}_{i,k|k-1}^{(1)} - \hat{\boldsymbol{x}}_{i,k|k-1}), \ldots, (\hat{\boldsymbol{x}}_{i,k|k-1}^{(N_e)} - \hat{\boldsymbol{x}}_{i,k|k-1})]$. The cross covariance between $\boldsymbol{x}_{i,k}$ and $\boldsymbol{y}_{i,k}$, given $\mathbb{Y}_{i,k-1}$, can be approximated as

[1] M/N must be an integer.

$$\boldsymbol{P}^{xy}_{i,k|k-1} = \frac{1}{N_e - 1}\boldsymbol{E}^{x}_{i,k|k-1}\left(\boldsymbol{E}^{y}_{i,k|k-1}\right)^{\mathrm{T}}. \tag{13}$$

At the ith local filter, the conditional probability distribution $p(\boldsymbol{x}_{i,k}|\mathbb{Y}_{i,k})$ is approximately represented via the *a posteriori* ensemble associated with the kth time slot, denoted $\{\hat{\boldsymbol{x}}^{(j)}_{i,k|k}, 1 \leq j \leq N_e\}$ and computed, exploiting the latest measurement $\boldsymbol{y}_{i,k}$, according to the following updating mechanism:

$$\hat{\boldsymbol{x}}^{(j)}_{i,k|k} = \hat{\boldsymbol{x}}^{(j)}_{i,k|k-1} + \boldsymbol{K}_{i,k}\left(\boldsymbol{y}_{i,k} - \hat{\boldsymbol{y}}^{(j)}_{i,k|k-1}\right) \tag{14}$$

$$\boldsymbol{K}_{i,k} = \boldsymbol{P}^{xy}_{i,k|k-1}\left(\boldsymbol{P}^{y}_{i,k|k-1}\right)^{-1}. \tag{15}$$

Finally, the updated estimate of the mean and covariance are

$$\hat{\boldsymbol{x}}_{i,k|k} = \frac{1}{N_e}\sum_{j=1}^{N_e}\hat{\boldsymbol{x}}^{(j)}_{i,k|k} \tag{16}$$

$$\boldsymbol{P}_{i,k|k} = \frac{1}{N_e - 1}\boldsymbol{E}^{x}_{i,k|k}\left(\boldsymbol{E}^{x}_{i,k|k}\right)^{\mathrm{T}} \tag{17}$$

where $\boldsymbol{E}^{x}_{i,k|k} = [(\hat{\boldsymbol{x}}^{(1)}_{i,k|k} - \hat{\boldsymbol{x}}_{i,k|k}), \ldots, (\hat{\boldsymbol{x}}^{(N_e)}_{i,k|k} - \hat{\boldsymbol{x}}_{i,k|k})]$.

It is worth mentioning that the grouping technique outlined in Section III-A is pivotal in fault detection and isolation as it ensures that each local filter can independently generate its estimates using a distinct subset of measurements. The generic local filter computes the local state vector estimate using the local subset of measurements, which exclusively impacts the corresponding elements in the local state estimate, leaving the other local filters unaffected. Additionally, the Kalman gain matrix ($\boldsymbol{K}_{i,k}$) significantly influences the *a posteriori* state vector estimate ($\hat{\boldsymbol{x}}_{i,k|k}$) of the local filter [23], [55], maintaining nonzero values for elements associated with the local measurements ($\boldsymbol{y}_{i,k}$) and setting the rest to zero. For example, if the cth measurement is absent from the local subset, all the elements in the cth column of the gain matrix are set to zero, and then the *a posteriori* estimates of the remaining elements remain equal to the *a priori* estimates. In the subsequent section, we will delve into how this grouping technique facilitates fault detection and isolation.

4) Information Fusion in 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 computed as [32]

$$\boldsymbol{P}_{k|k} = \left(\sum_{i=1}^{N}\boldsymbol{P}^{-1}_{i,k|k}\right)^{-1},\ \hat{\boldsymbol{x}}_{k|k} = \boldsymbol{P}_{k|k}\sum_{i=1}^{N}\boldsymbol{P}^{-1}_{i,k|k}\hat{\boldsymbol{x}}_{i,k|k}. \tag{18}$$

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

All the steps involved in the state estimation procedure (related to computations at both main and local filters) are outlined in Algorithm 1.

5) Fault Detection and Isolation: For fault diagnosis, we employ the local state residual vector ($\boldsymbol{r}_{i,k} \in \mathbb{R}^{n_x \times 1}$) as decision statistic, whose ℓth entry is defined as

Algorithm 1: Implementation of the Partial-Distributed State Estimation During the kth Time Instant

Input : *A posteriori* state estimate $\hat{\boldsymbol{x}}_{k-1|k-1}$ during the $(k-1)$th time instant, number of local filters N and ensemble size N_e

Main filter:

  Generate ensemble of samples $\{\hat{\boldsymbol{x}}^{(j)}_{k-1|k-1}\}$**:**

    for each sample $j = 1, \ldots, N_e$ **do**

      State vector ensemble generation via Eq. (6)

    end for

  Generate *a priori* ensemble $\{\hat{\boldsymbol{x}}^{(j)}_{k|k-1}\}$**:**

    for each sample $j = 1, \ldots, N_e$ **do**

      Process model update via Eq. (7)

    end for

  N **local filters:**

    for each local filter $i = 1, \ldots, N$ **do**

      Assign *a priori* ensemble to *i*th local:

        for each sample $j = 1, \ldots, N_e$ **do**

          $\hat{\boldsymbol{x}}^{(j)}_{i,k|k-1} = \hat{\boldsymbol{x}}^{(j)}_{k|k-1}$

        end for

      Generate local measurement ensemble $\{\hat{\boldsymbol{y}}^{(j)}_{i,k|k-1}\}$**:**

        for each sample $j = 1, \ldots, N_e$ **do**

          Measurement model update via Eq. (8)

        end for

      Generate *a priori* local state estimate $\hat{\boldsymbol{x}}_{i,k|k-1}$ **and local covariance matrix** $\boldsymbol{P}_{i,k|k-1}$**:**

        Sample mean and covariance of the measurement ensemble via Eqs. (9) and (10)

        Sample mean and covariance of the *a priori* ensemble via Eqs. (11) and (12)

      Generate *a posteriori* ensemble $\{\hat{\boldsymbol{x}}^{(j)}_{i,k|k}\}$ **for *i*th local:**

        for each sample $j = 1, \ldots, N_e$ **do**

          Update each *a priori* ensemble sample by incorporating the latest measurement via Eqs. (14) and (15)

        end for

      Compute local state estimate $\hat{\boldsymbol{x}}_{i,k|k}$ **and local covariance matrix** $\boldsymbol{P}_{i,k|k}$**:**

        Obtain sample mean and covariance of the *a posteriori* ensemble via Eqs. (16) and (17)

    end for

Main filter:

  Combine N local filter outputs in the information fusion to generate global outputs using Eq. (18)

Output: Global state estimate $\hat{\boldsymbol{x}}_{k|k}$, and global state covariance matrix $\boldsymbol{P}_{k|k}$ at the kth time instant

$$r_{i,k;\ell} = \left|\hat{x}_{i,k|k;\ell} - \frac{1}{N}\sum_{j=1}^{N}\left(\hat{x}_{j,k|k;\ell}\right)\right| \tag{19}$$

being $\hat{x}_{i,k|k;\ell}$ the ℓth entry of the state vector estimate $\hat{\boldsymbol{x}}_{i,k|k}$.

The local state residual vector quantifies the deviation among the independent local state vector estimates. The local state estimate exhibits high residual values only when its associated sensor measurements are faulty. The evaluation based on the residual values is directly linked with the sensor grouping discussed in Section III-A. Each local filter processes a distinct subset of sensor measurements, updating only part of the local state vector estimate. When a sensor fault occurs, it affects only the corresponding elements of the state estimate of the local filter with the faulty measurements, while other filters' estimates remain unaffected. The residual, measuring the deviation of one local filter estimate from the mean of all local filters' estimates, becomes high for faulty estimates. For example, if the M/Nth measurement of local filter 1 is faulty, it only impacts the M/Nth element in the state estimate of the local filter 1 and M/Nth estimate of remaining $N-1$ local filters remain nonfaulty. Further, the mean M/Nth state estimate remains close to the nonfaulty value since most local filters provide nonfaulty estimates, resulting in a high residual for the 1st local filter and low residuals for other $N-1$ local filters.

By comparing the local state residual entries against predefined thresholds, the presence of faults within the system can be easily identified: if the ℓth state residual entry $r_{i,k:\ell}$ exceeds the threshold γ_ℓ, a faulty state vector estimate at the ℓth position of the ith local is declared, and its faulty local state estimate $\hat{\boldsymbol{x}}_{i,k|k}$ can be corrected as

$$\hat{x}_{i,k|k;\ell} = \begin{cases} \hat{x}_{i,k|k;\ell}, & r_{i,k;\ell} < \gamma_\ell \\ \hat{x}_{i,k|k-1;\ell}, & r_{i,k;\ell} > \gamma_\ell \end{cases} \tag{20}$$

where the ℓth faulty entry of the state estimate $\hat{x}_{i,k|k;\ell}$ is substituted with the ℓth entry of the *a priori* state estimate $\hat{x}_{i,k|k-1;\ell}$ corresponding to the ith local filter, which is not affected by the faulty local measurement. The use of *a priori* state estimate ensures that no information is lost during fault isolation.

It is worth highlighting that traditional fault detection methods often rely on fixed thresholding techniques, which may face challenges in adapting to dynamic environmental conditions, resulting in excessive false positives or missed detections. Therefore, to enhance the accuracy and responsiveness of the fault detection system, it becomes essential to employ adaptive thresholding methods for fault diagnosis. We propose an adaptive threshold mechanism based on the local state residual detection metric. Since the local state residual becomes higher for a faulty local and lower for the unaffected ones, the threshold for identifying a faulty local is chosen to be higher than the average value of the local state residuals across all local filters, i.e.,

$$\gamma_{\ell,k} = \frac{1}{N}\sum_{i=1}^{N} r_{i,k;\ell} + \lambda \tag{21}$$

where the tuning factor λ guarantees a nonzero and higher residual threshold in contrast to the average value. Notably, the threshold must remain below the residual of a faulty estimate and higher than those of nonfaulty estimates. Moreover, the adaptive threshold only depends on the tuning factor λ and is independent of the process and measurement noise amplitudes. Following the identification of faulty estimates, the local state estimates are combined using the information fusion.

Algorithm 2: Proposed SFDIA Architecture in the Presence of Multiple Sensor Faults

Input : Number of local filters N, boundary condition $\boldsymbol{u}_{bc}$ and ensemble size N_e
Initialization: $\hat{\boldsymbol{x}}_{0|0} = \hat{\boldsymbol{x}}_0$,
while $k < k_{end}$ **do**
  Main filter: Compute *a priori* ensemble $\{\hat{\boldsymbol{x}}^{(j)}_{k|k-1}\}$
  Local filters: Compute local estimates $\hat{\boldsymbol{x}}_{i,k|k}$ and $\boldsymbol{P}_{i,k|k}$ using linear measurement updates
  Fault detection:
    Compute state-residual vector $\boldsymbol{r}_{i,k}$ and adaptive threshold $\boldsymbol{r}_{th,k}$ via Eqs. (19) and (21)
    if $r^{(\ell)}_{i,k} > r^{(\ell)}_{th,k}$ **then**
      Fault detected
      Local state estimate correction based on residual via Eq. (20)
    else
      No fault
    end if
  Information fusion: Compute global estimates $\hat{\boldsymbol{x}}_{k|k}$ and $\boldsymbol{P}_{k|k}$ using fault-free local state estimates
  $k \leftarrow k+1$
end while
Output: Fault-free global estimates $\hat{\boldsymbol{x}}_{k|k}$ and $\boldsymbol{P}_{k|k}$

The various steps involved in implementing the proposed SFDIA architecture are summarized in Algorithm 2.

IV. Numerical Results

The effectiveness of our proposed architecture is assessed using simulated data generated according to the transient flow model discussed in Section II and with reference to a high-pressure hydrogen-blended natural gas pipeline. The relevant parameters are specified in Table II. We assume that the natural gas is blended with 10%-volume hydrogen.[2] MATLAB software is utilized to derive the numerical results. Experiments were developed (and here described accordingly) in two parts[3]: the former focusing on the state-estimation capabilities and the latter focusing on the fault-diagnosis capabilities.

The boundary conditions are set as $p(0,t) = 8.4$ MPa, $T(0,t) = 290.15$ K, and $\dot{m}(L,t) = f(t)$ (as shown in Fig. 2(a). The simulations span a time period of $t_f \in [0, 7200$ s], while the spatial and temporal step sizes are $\Delta s = 3540$ m and $\Delta t = 5$ s, respectively. The corresponding simulated data without any observation noise are shown in Fig. 2(b)–(d). Additionally, Fig. 2(a) illustrates the flow rate boundary conditions, emphasizing swift variations within the

[2]To avoid modifications in gas pipelines, that are in service for many years, the limits set for hydrogen by manufacturers are 5% and 10% [56].

[3]100 iterations of Monte Carlo Runs are conducted for the experiments.

TABLE II
PARAMETERS CONSIDERED DURING SIMULATIONS

Parameters	Values
L	177 km
d	1.3796 m
ϵ	0.01922 mm
T_s	12°C
U	2.0 $\text{Wm}^{-2}\text{K}^{-1}$

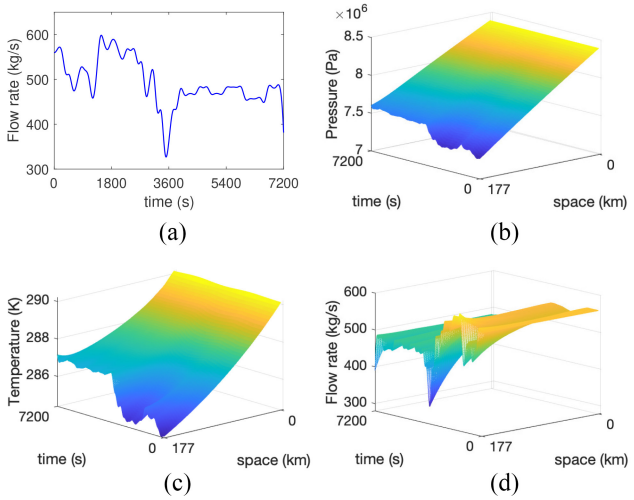

Fig. 2. Simulated data using ODEs. (a) Boundary condition. (b) Pressure. (c) Temperature. (d) Flow rate.

solution domain. These variations may arise from disruptions in gas demand or shifts in the operational settings of control devices like valves, compressors, and pressure regulators.

The noisy fault-free measurement ($y_j[k]$) for the jth sensor at the kth time instant is produced by adding zero-mean white Gaussian noise ($q_j[k]$) to the originally generated noiseless and fault-free value ($x_j[k]$), i.e.,

$$y_j[k] = x_j[k] + q_j[k]. \tag{22}$$

To validate the robustness of the state estimation by the proposed architecture, we generated measurement noise considering three levels of signal-to-noise ratio (SNR): 1) low; 2) moderate; and 3) high. The distribution of the measurement noise for each of these scenarios is illustrated in Table III.

In our experiments, a total of $N = 153$ sensors is considered, consisting of 51 pressure sensors, 51 flow-rate sensors, and 51 temperature sensors. We choose the number of local filters as $M = 3$, where each local filter receives an array of 51 sensor measurements of the same type. Within the partial-distributed EnKF framework, an ensemble size of $N_e = 100$ is chosen to optimize the performance. Additionally, the standard deviation of the process noise is adjusted to be 10% lower than that of the measurement noise. The matrices $\boldsymbol{Q}_{i,k}$ and $\boldsymbol{R}_{i,k}$ are considered as diagonal matrices, where the diagonal elements represent the variances of the process noise and measurement noise, respectively.

The estimation performance is evaluated in terms of the root mean square error (RMSE) in both spatial and temporal

TABLE III
MEASUREMENT NOISE ASSOCIATED WITH PRESSURE, TEMPERATURE, AND FLOW RATE FOR VARIOUS SNR SCENARIOS

Measurement Noise	Pressure (10^{-3} MPa)	Flow Rate (kg/s)	Temperature (K)
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 IV
COMPARING THE RMSE OF DIFFERENT FILTERS IN NONFAULTY SCENARIOS

SNR	Filter	Pressure (10^{-3} MPa)	Flow Rate (kg/s)	Temperature (K)
High	Proposed PD	0.180	1.8225	0.1006
	Proposed FD	0.245	1.2665	0.8833
	MM-SFDIA	0.4325	1.1726	1.7361
	Conventional EnKF	**0.0553**	**0.1987**	**0.0966**
Moderate	Proposed PD	**0.07532**	2.0996	**0.0366**
	Proposed FD	0.800	1.9316	1.6145
	MM-SFDIA	0.9143	2.7952	4.1291
	Conventional EnKF	0.12903	**0.7835**	0.3121
Low	Proposed PD	**0.126**	**1.2294**	**0.0541**
	Proposed FD	1.8814	5.6385	3.8561
	MM-SFDIA	2.3472	3.2877	2.1509
	Conventional EnKF	11.640	3.4071	1.0599

dimensions, which is averaged over τ iterations, ensuring distinct random seeds for each iteration, i.e.,

$$\text{RMSE} = \frac{1}{\tau}\sum_{i=1}^{\tau}\left(\frac{\|\boldsymbol{X} - \hat{\boldsymbol{X}}\|_F}{\sqrt{nN_k}}\right)$$

where $\|\cdot\|_F$ represents the Frobenius norm, and the true state matrix $\boldsymbol{X}$ and the estimated state matrix $\hat{\boldsymbol{X}}$ are defined as $\boldsymbol{X} = [\boldsymbol{x}_{i,1|1}, \boldsymbol{x}_{i,2|2}, \ldots, \boldsymbol{x}_{i,N_k|N_k}]$ and $\hat{\boldsymbol{X}} = [\hat{\boldsymbol{x}}_{i,1|1}, \hat{\boldsymbol{x}}_{i,2|2}, \ldots, \hat{\boldsymbol{x}}_{i,N_k|N_k}]$, respectively. Additionally, the number of spatial nodes n and the number of time steps N_k can be evaluated using the expressions $n = L/\Delta s$ and $N_k = t_f/\Delta t$, respectively. The estimation performance of the proposed partial distributed (proposed PD) technique discussed in Section III is compared against different baselines, including the fully distributed version of the proposed architecture (proposed FD); the model-based multisensor fault detection, isolation and accommodation (MM-SFDIA) discussed in [40], utilizing a centralized EnKF-based technique for state estimation; and the centralized conventional EnKF [23]. In the Proposed FD version, both nonlinear time updates and linear measurement updates are carried out by the local filters, while the main filter only performs the information fusion. The RMSE results are explicitly presented in Table III, and it is apparent that the proposed PD achieves a comparable level of estimation accuracy when compared to the fully distributed version and to other baselines especially in case of low SNR. Compared to MM-SFDIA and the centralized conventional EnKF, the proposed PD uses a PD architecture that utilizes a specific grouping of measurements discussed in Section III-A. This grouping introduces numerous redundant state estimates that are not affected by measurement noise, making the state estimation of the proposed PD more robust and comparable.

TABLE V
COMPUTATIONAL COMPLEXITY

Technique	Computation complexity	
	Main Filter	ith Local filter
Proposed PD	$N_e(n_x^2+3n_x)+2Nn_x^2+2n_x^3$	$N(N_e(n_y^2+3n_y+n_x^2+2n_x)+2n_xn_y+n_y^3)$
Proposed FD	$2Nn_x^2+2n_x^3$	$N(N_e(n_y^2+3n_y+2n_x^2+5n_x)+2n_xn_y+n_y^3)$
Conventional EnKF	$N_e(6n_x^2+9n_x)+2n_x^3$	–

Moreover, the computational complexity of our proposed PD technique is compared against baselines, including the proposed FD and conventional EnKF, as shown in Table V. It is clearly demonstrated that our technique has low complexity compared to the proposed FD and centralized approach for $n_y < n_x$. Consequently, our proposed PD method achieves better RMSE with reduced computational complexity compared to the baselines.

To assess the performance of the proposed estimation method with SFDIA mechanism, synthetically generated fault signals are introduced into the simulated data obtained from the transient-flow model. It is worth noting that sensors can encounter various types of faults, including bias, drift, freezing, and random faults, which are commonly observed issues, as discussed in [9] and [10]. Without compromising a broad perspective of generality, here bias and drift faults are specifically employed to demonstrate hard and soft failures, respectively. The mathematical expressions of these fault types are discussed as follows.

Bias Fault: A constant offset, denoted by bias (b), is added to the sensor measurements for a period of G consecutive samples, which is expressed as

$$\tilde{y}_j[k] = \begin{cases} y_j[k]+b, & 0 \le k-g \le G \\ y_j[k], & \text{else} \end{cases}$$

where $\tilde{y}_j[k]$ represents the faulty measurement, and g indicates the time when the fault begins.

Drift Fault: The actual measurement gradually increases (up to the maximum bias level b) for G time instances, as

$$\tilde{y}_j[k] = \begin{cases} y_j[k]+\frac{b(k-g+1)}{G}, & 0 \le k-g \le G \\ y_j[k]+b, & G \le k-g \le G+K \\ y_j[k], & \text{else} \end{cases}$$

where the variable K represents the number of samples during which the drift fault maintains the saturated bias level b. Additionally, we highlight the impact of the drift fault by considering $G > K$.

The efficacy of our proposed architecture is assessed in the presence of weak and strong faults for both types (bias and drift). An absolute level b within a uniform distribution spanning 20%–40% (resp. 60%–90%) of the data amplitude is considered for the weak (resp. strong) faults.[4] The fault duration is set by randomly selecting $G \in \{5, 6\}$ and $K \in \{3, 4\}$ consecutive samples. The randomness in selecting the level and the duration of the faults is meant to test the robustness to arbitrary fault characteristics.

[4]Both positive and negative faults are employed in a randomized manner to enhance robustness.

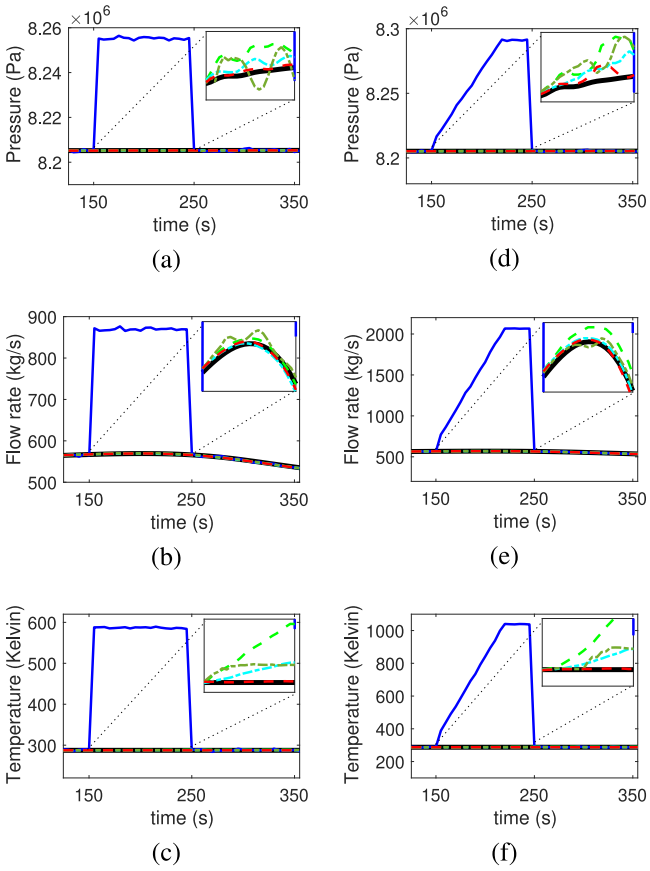

Fig. 3. State estimation techniques comparison (classic EnKF (cyan), MM-SFDIA (dark green) [40], proposed PD without SFDIA (green), and proposed PD with SFDIA (red) in the presence of three simultaneous bias and drift faults occurring at indices $\ell = 10, 100, 121$. Bias and drift fault results are depicted in plots (a)–(c) and plots (d)–(f), respectively. Actual values are indicated in black, while faulty values are shown in blue.

Fig. 3 shows how the various algorithms behave in presence of multiple faults, with reference to 3 simultaneous bias and drift faults located in the interval (150 s, 250 s) introduced to the pressure, flow rate, and temperature sensors indexed with $\ell = 10, 100$, and 121. More specifically, the comparison of the actual state and the estimated state values obtained using different approaches (our novel architecture with and without SFDIA, MM-SFDIA integrating multiple fault detection, and the classic EnKF method without SFDIA) illustrates the superiority of our proposed architecture with SFDIA.

The detection performance of the proposed architecture is further assessed in Fig. 4 in terms of the receiver operating characteristic (ROC), contrasting the probabilities of detection and false alarm calculated on a per-sample basis. These plots

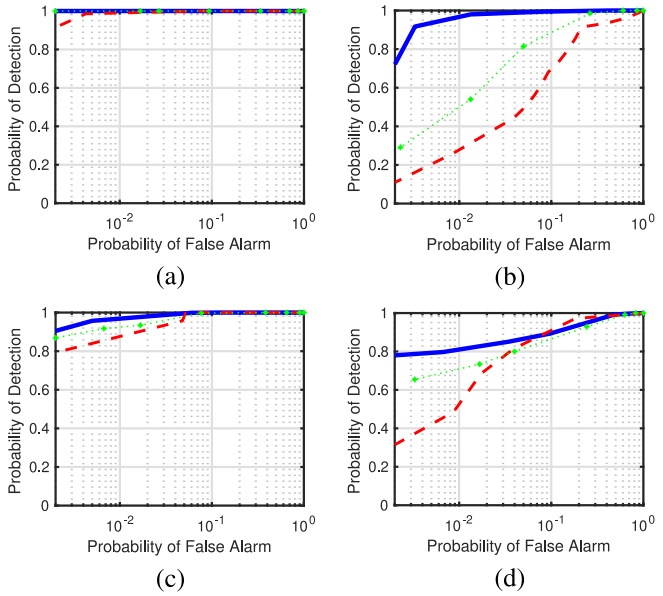

Fig. 4. ROC plots indicating the detection performance of proposed PD (blue), proposed FD (green), and MM-SFDIA (red) for three simultaneous faults of different types. (a) Strong bias. (b) Weak bias. (c) Strong drift. (d) Weak drift.

TABLE VI
COMPARING THE ACCURACY OF THE PROPOSED PD DESIGN WITH THE PROPOSED FD, AND MM-SFDIA WHILE HANDLING MULTIPLE SENSOR FAULTS

Method	High Bias	Low Bias	High Drift	Low Drift
Proposed PD	**0.99**	**0.95**	**0.98**	**0.89**
Proposed FD	0.96	0.88	0.96	0.87
MM-SFDIA	0.97	0.84	0.95	0.88

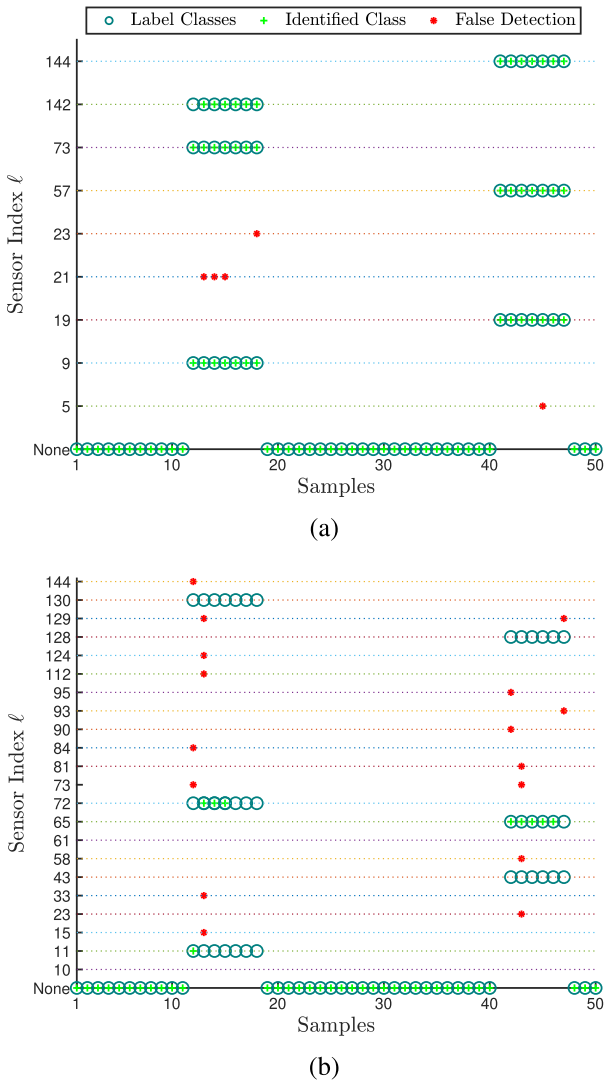

Fig. 5. Fault isolation and detection visualization of (a) proposed PD architecture and (b) MM-SFDIA.

are generated by varying the tuning parameter λ of the adaptive threshold. The result confirms that our proposed PD technique attains a significantly higher detection probability and considerably lower false alarm probability, particularly in the weak fault scenario, when compared to MM-SFDIA, a recently proposed method explicitly designed for addressing multiple sensor faults. Moreover, a comparison among the architectures mentioned above is also demonstrated in Table VI in terms of accuracy.

The robustness of the architecture was assessed with respect to the number of faulty sensors, with respect to the number of local filters, and with respect to the bias level. Numerical simulations showed that the system accuracy remains unaffected for scenarios with up to 9 faulty sensors, up to 21 local filters and bias level down to 3%.

Moreover, the fault isolation capability of the proposed PD method and MM-SFDIA [40] in the presence of multiple faults is clearly illustrated in Fig. 5. A total of 50 samples is considered, with a focus on weak bias faults: the "○" symbol denotes the actual fault, while "None" signifies a scenario without any faults within the system. The falsely detected sensor fault is specified with the "∗" symbol and "+" symbol indicates the correctly identified fault. It can be observed that the proposed architecture can effectively detect all sensor faults with only a few instances of false detection. On the contrary, MM-SFDIA fails to correctly identify and isolate the faulty sensors.

Our proposed PD approach, exploiting the partial-distributed state estimation framework, does not compromise the state estimation accuracy while significantly reducing the computational requirements as illustrated in Table VII in terms of the average execution time and standard deviation. These results demonstrate the efficacy of our proposed architecture achieving high accuracy with low complexity in the presence of multiple faults in a nonlinear highly complex system dealing with hydrogen-blended natural gas.

V. CONCLUSION

In this article, a novel approach for state estimation and sensor fault diagnosis in hydrogen-blended natural gas pipelines experiencing transient flow is introduced. Our method incorporates a unique architecture based on a partial-distributed EnKF framework. The proposed design effectively segregates

TABLE VII
COMPARING THE EXECUTION TIME OF THE PROPOSED PD ARCHITECTURE WITH THE PROPOSED FD, MM-SFDIA, AND CONVENTIONAL ENKF FOR HANDLING MULTIPLE SENSOR FAULTS

Method	Average Execution Time	Standard deviation
Proposed PD	**0.27** s	1.3e-05 s
Proposed FD	0.78 s	1.1e-03 s
MM-SFDIA	0.68 s	1.4e-04 s
Conventional EnKF	0.44 s	9.7e-05 s

the nonlinear computations from the local filters and redirects them to the main filter, thereby significantly reducing the computational overhead. Additionally, a new fault detection and isolation technique, utilizing the local state residuals, is developed to address simultaneous multiple sensor faults. Furthermore, an adaptive thresholding technique has been implemented to enhance fault detection accuracy. Through extensive simulations, we showcased the effectiveness of our proposed architecture in accurately detecting and isolating multiple sensor faults occurring simultaneously. Future work aims to explore: 1) identifying both process-related issues (such as leak detection) and sensor faults in natural gas pipelines; 2) detecting faults in pipelines experiencing varying operating conditions; and 3) designing control laws or observers for PDEs governing the flow model.

REFERENCES

[1] L. Hou, D. Wang, B. Du, X. Qian, and M. Yuan, "Gas concentration detection via multi-channeled air sampling method," *Sens. Rev.*, 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," *J. Safety Environ.*, vol. 19, no. 3, pp. 902–908, 2019.

[3] T. B. Quy and J.-M. Kim, "Real-time leak detection for a gas pipeline using a k-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, Nov. 2019.

[6] A. Chawla et al., "IoT-based monitoring in carbon capture and storage systems," *IEEE Internet Things Mag.*, vol. 5, no. 4, pp. 106–111, Dec. 2022.

[7] Y. Feng and H.-X. Li, "Detection and spatial identification of fault for parabolic distributed parameter systems," *IEEE Trans. Ind. Electron.*, vol. 66, no. 9, pp. 7300–7309, Sep. 2019.

[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," *Int. J. Control, Autom. Syst.*, vol. 13, no. 5, pp. 1037–1046, 2015.

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

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

[11] M. Zhao and O. Fink, "DyEdgeGAT: Dynamic edge via graph attention for early fault detection in IIoT systems," *IEEE Internet Things J.*, vol. 11, no. 13, pp. 22950–22965, Jul. 2024.

[12] S. Lu, J. Lu, K. An, X. Wang, and Q. He, "Edge computing on IoT for machine signal processing and fault diagnosis: A review," *IEEE Internet Things J.*, vol. 10, no. 13, pp. 11093–11116, Jul. 2023.

[13] 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," *Inf. Sci.*, vol. 259, pp. 304–320, Feb. 2014.

[14] Q. Zhang, "Dynamic system fault diagnosis under sparseness assumption," *IEEE Trans. Signal Process.*, vol. 69, pp. 2499–2508, Apr. 2021.

[15] J. Zhao et al., "Review of lithium-ion battery fault features, diagnosis methods and diagnosis procedures," *IEEE Internet Things J.*, vol. 11, no. 11, pp. 18936–18950, Jun. 2024.

[16] M. Zhong, X. Du, Y. Song, T. Xue, and S. X. Ding, "Event-triggered parity space approach to fault detection for linear discrete-time systems," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 52, no. 8, pp. 4813–4822, Aug. 2022.

[17] B. Pourbabaee, N. Meskin, and K. Khorasani, "Robust sensor fault detection and isolation of gas turbine engines subjected to time-varying parameter uncertainties," *Mech. Syst. Signal Process.*, vol. 76, pp. 136–156, Aug. 2016.

[18] J. Zarei and E. Shokri, "Robust sensor fault detection based on nonlinear unknown input observer," *Measurement*, vol. 48, pp. 355–367, Feb. 2014.

[19] G. Tabella, D. Ciuonzo, N. Paltrinieri, and P. S. Rossi, "Bayesian fault detection and localization through wireless sensor networks in industrial plants," *IEEE Internet Things J.*, vol. 11, no. 8, pp. 13231–13246, Apr. 2024.

[20] X. Wei, M. Verhaegen, and T. van Engelen, "Sensor fault detection and isolation for wind turbines based on subspace identification and Kalman filter techniques," *Int. J. Adapt. Control Signal Process.*, vol. 24, no. 8, pp. 687–707, 2010.

[21] J. Wei, G. Dong, and Z. Chen, "Model-based fault diagnosis of lithium-ion battery using strong tracking extended Kalman filter," *Energy Procedia*, vol. 158, pp. 2500–2505, Feb. 2019.

[22] W. El Sayed, M. A. 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.

[23] H. Fang, N. Tian, Y. Wang, M. Zhou, and M. A. Haile, "Nonlinear Bayesian estimation: From Kalman filtering to a broader horizon," *IEEE/CAA J. Automatica Sinica*, vol. 5, no. 2, pp. 401–417, Mar. 2018.

[24] L. Yan, Y. Zhang, B. Xiao, Y. Xia, and M. Fu, "Fault detection for nonlinear systems with unreliable measurements based on hierarchy cubature Kalman filter," *Can. J. Chem. Eng.*, vol. 96, no. 2, pp. 497–506, 2018.

[25] 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," *Meas. Sci. Technol.*, vol. 32, no. 9, 2021, Art. no. 95905.

[26] D. Mori, H. Sugiura, and Y. Hattori, "Adaptive sensor fault detection and isolation using unscented Kalman filter for vehicle positioning," in *Proc. IEEE Intell. Transp. Syst. Conf. (ITSC)*, 2019, pp. 1298–1304.

[27] 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.

[28] Y. Luo, Y. Liu, W. Yang, J. Zhou, and T. Lv, "Distributed filtering algorithm based on local outlier factor under data integrity attacks," *J. Frankl. Inst.*, vol. 360, no. 12, pp. 9290–9306, 2023.

[29] N. S. Nokhodberiz and J. Poshtan, "Belief consensus–based distributed particle filters for fault diagnosis of non-linear distributed systems," *Proc. Inst. Mech. Eng., Part I, J. Syst. Control Eng.*, vol. 228, no. 3, pp. 123–137, 2014.

[30] 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 Dyn.*, vol. 111, no. 3, pp. 2571–2589, 2023.

[31] K. Geng and N. A. Chulin, "Applications of multi-height sensors data fusion and fault-tolerant Kalman filter in integrated navigation system of UAV," *Procedia Comput. Sci.*, vol. 103, pp. 231–238, Jan. 2017.

[32] A. Vafamand, B. Moshiri, and N. Vafamand, "Fusing unscented Kalman filter to detect and isolate sensor faults in DC microgrids with CPLs," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–8, Dec. 2022.

[33] K. Salahshoor, M. Mosallaei, and M. Bayat, "Centralized and decentralized process and sensor fault monitoring using data fusion based on adaptive extended Kalman filter algorithm," *Measurement*, vol. 41, no. 10, pp. 1059–1076, 2008.

[34] Z. Xing and Y. Xia, "Distributed federated Kalman filter fusion over multi-sensor unreliable networked systems," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 63, no. 10, pp. 1714–1725, Oct. 2016.

[35] K. Bader, B. Lussier, and W. Schön, "A fault tolerant architecture for data fusion: A real application of Kalman filters for mobile robot localization," *Robot. Auton. Syst.*, vol. 88, pp. 11–23, Feb. 2017.

[36] H. Hamadi, B. Lussier, I. Fantoni, and C. Francis, "Data fusion fault tolerant strategy for a quadrotor UAV under sensors and software faults," *ISA Trans.*, vol. 129, pp. 520–539, Oct. 2022.

[37] M. Saied, A. R. Tabikh, C. Francis, H. Hamadi, and B. Lussier, "An informational approach for fault tolerant data fusion applied to a UAV's attitude, altitude, and position estimation," *IEEE Sensors J.*, vol. 21, no. 24, pp. 27766–27778, Dec. 2021.

[38] D.-J. Xin, L.-F. Shi, and X. Yu, "Distributed Kalman filter with faulty/reliable sensors based on Wasserstein average consensus," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 69, no. 4, pp. 2371–2375, Apr. 2022.

[39] C. Liang, F. Wen, and Z. Wang, "Trust-based distributed Kalman filtering for target tracking under malicious cyber attacks," *Inf. Fusion*, vol. 46, pp. 44–50, Mar. 2019.

[40] M. Doostmohammadian and N. Meskin, "Sensor fault detection and isolation via networked estimation: Full-rank dynamical systems," *IEEE Trans. Control Netw. Syst.*, vol. 8, no. 2, pp. 987–996, Jun. 2021.

[41] F. Lu, Z. Li, J. Huang, and M. Jia, "Hybrid state estimation for aircraft engine anomaly detection and fault accommodation," *AIAA J.*, vol. 58, no. 4, pp. 1748–1762, 2020.

[42] 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 *Proc. IEEE SENSORS*, 2023, pp. 1–4.

[43] k. Shaheen, A. Chawla, F. E. Uilhoorn, and P. S. Rossi, "Sensor-fault detection, isolation and accommodation for natural-gas pipelines under transient flow," *IEEE Trans. Signal Inf. Process. Over Netw.*, vol. 10, no. 2, pp. 264–276, Mar. 2024.

[44] F. Lu, Y. Wang, J. Huang, Y. Huang, and X. Qiu, "Fusing unscented Kalman filter for performance monitoring and fault accommodation in gas turbine," *Proc. Inst. Mech. Eng., Part G, J. Aerosp. Eng.*, vol. 232, no. 3, pp. 556–570, 2018.

[45] 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 J.*, vol. 24, no. 3, pp. 3554–3567, Feb. 2024.

[46] F. Uilhoorn and M. Witek, "Influence of hydrogen blended gas transmission under transient flow on L485ME steel grade fracture toughness," *Gaz, Woda i Technika Sanitarna*, vol. 44, no. 12, pp. 1–9, 2023.

[47] F. E. Uilhoorn, "Comparison of Bayesian estimation methods for modeling flow transients in gas pipelines," *J. Nat. Gas Sci. Eng.*, vol. 38, pp. 159–170, Feb. 2017.

[48] 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," *Comput. Math. Appl.*, vol. 68, no. 12, pp. 1991–2004, 2014.

[49] F. E. Uilhoorn, "Dynamic behaviour of non-isothermal compressible natural gases mixed with hydrogen in pipelines," *Int. J. Hydrogen Energy*, vol. 34, no. 16, pp. 6722–6729, 2009.

[50] C. F. Colebrook et al., "Correspondence. turbulent flow in pipes, with particular reference to the transition region between the smooth and rough pipe laws.(includes plates)." *J. Inst. Civil Eng.*, vol. 12, no. 8, pp. 393–422, 1939.

[51] E. W. Lemmon and R. T. Jacobsen, "Viscosity and thermal conductivity equations for nitrogen, oxygen, argon, and air," *Int. J. Thermophysics*, vol. 25, pp. 21–69, Jan. 2004.

[52] W. E. Schiesser, *The Numerical Method of Lines: Integration of Partial Differential Equations*. Elsevier, Amsterdam, The Netherlands, 2012.

[53] R. Courant, "Uber die partiellen Differenzengleichungen der mathematischen Physik," *Math. Ann.*, vol. 100, no. 1, pp. 32–74, 1928.

[54] R. Olfati-Saber, "Distributed Kalman filtering for sensor networks," in *Proc. 46th IEEE Conf. Decis. Control*, 2007, pp. 5492–5498.

[55] D. Assimilation, "The ensemble Kalman filter," *IEEE Control Syst. Mag.*, to be published.

[56] M. Witek and F. Uilhoorn, "Impact of hydrogen blended natural gas on linepack energy for existing high pressure pipelines," *Arch. Thermodyn.*, vol. 43, no. 3, pp. 111–124, 2022.

Khadija Shaheen (Member, IEEE) received the bachelor's degree in electrical engineering from Riphah International University, Islamabad, Pakistan, in 2017, and the M.Sc. degree in electrical engineering with specialization in digital systems and signal processing from the School of Electrical Engineering and Computer Science, National University of Sciences and Technology, Islamabad, in 2021. She is currently pursuing the Ph.D. degree in electronics with the Department of Electronic Systems, Norwegian University of Science and Technology, Trondheim, Norway.

Her research interests include sensor validation, machine learning, and digital signal processing.

Apoorva Chawla (Member, IEEE) received the B.Tech. degree in electronics and communication engineering from the Gautam Buddh Technical University, Greater Noida, India, in 2012, and the M.Tech. and Ph.D. dual degrees in electrical engineering from the Indian Institute of Technology Kanpur, Kanpur, India, in 2022.

She is currently a Postdoctoral Researcher with the Department of Electronic Systems, Norwegian University of Science and Technology, Trondheim, Norway. Her research interests include sensor validation, the Internet of Things, distributed detection, wireless communication, and signal processing.

Dr. Chawla was awarded the TCS Research Fellowship for Pursuing Graduate Studies at the Indian Institute of Technology Kanpur, Kanpur, India. In 2019, she was selected as One of the Finalists for the Qualcomm Innovation Fellowship by Qualcomm, India.

Ferdinand Evert Uilhoorn received the B.Sc. and M.Sc. degrees in mechanical and systems engineering from Polytechnic Groningen and Delft University of Technology, Delft, The Netherlands, and the Ph.D. (cum laude) and D.Sc. degrees in gas engineering from Warsaw University of Technology, Warsaw, Poland, in 2007 and 2017, respectively.

He is currently an Associate Professor with the Gas Engineering Group, Warsaw University of Technology. His research interests include single- and multi-phase flow modeling, numerical methods, and data assimilation, with a focus on pipeline systems.

Pierluigi Salvo Rossi (Senior Member, IEEE) was born in Naples, Italy, in 1977. He received the Dr.Eng. degree (summa cum laude) in telecommunications engineering and the Ph.D. degree in computer engineering from the University of Naples "Federico II," Naples, in 2002 and 2005, respectively.

He is currently a Full Professor and the Deputy Head with the Department of Electronic Systems, Norwegian University of Science and Technology (NTNU), Trondheim, Norway. He is also a Part-Time Research Scientist with the Department Gas Technology, SINTEF Energy Research, Trondheim. Previously, he was with the University of Naples "Federico II"; also with the Second University of Naples, Naples; also with NTNU; and also with Kongsberg Digital AS, Horten, Norway. He held visiting appointments with Drexel University, Philadelphia, PA, USA, Lund University, Lund, Sweden, NTNU, and Uppsala University, Uppsala, Sweden. His research interests fall within the areas of communication theory, data fusion, machine learning, and signal processing.

Prof. Salvo Rossi was awarded as an Exemplary Senior Editor of the IEEE Communications Letters in 2018. He is (or has been) in the Editorial Board of the IEEE SENSORS JOURNAL, IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY, the IEEE TRANSACTIONS ON SIGNAL AND INFORMATION PROCESSING OVER NETWORKS, the IEEE COMMUNICATIONS LETTERS, and the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.