12th IFAC International Symposium on
Advanced Control of Chemical Processes
July 14-17, 2024. Toronto, Canada

A distributed Kalman filter based fault detection scheme incorporating weighted average consensus algorithm for large-scale interconnected systems ⋆

Haozhou Yuan * **Jingjing Gao** * **Jian Huang** * **Xu Yang** *

* *Key Laboratory of Knowledge Automation for Industrial Processes of Ministry of Education, School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China (e-mail: M202210573@xs.ustb.edu.cn, jingjing.gao@ustb.edu.cn, jianhuang@ustb.edu.cn, yangxu@ustb.edu.cn)*

Abstract: Stimulated by the increasing demands for system safety and process reliability in complex industrial processes, this paper investigates a weighted average consensus algorithm based distributed fault detection scheme for large-scale interconnected systems, using a sensor network where each node is equipped with a Kalman filter (KF). To reduce the communication and computation efforts, the proposed distributed fault detection scheme is splitted into two phases: distributed offline training and online fault detection. To this end, the Expectation-Maximization (EM) algorithm is firstly addressed to identify the unknown measurement matrices and covariance matrices of noise vectors. It is followed by an average consensus algorithm so that the identical Kalman filters can be designed in parallel at all sensor nodes. On this basis, distributed residual generators and test statistics are constructed for fault detection purpose using the average consensus algorithm. Considering that there exist some special conditions, such as the occurrence of node failures, a variation of the distributed Kalman filter based fault detection scheme is proposed by dynamically adjusting the consensus weight. Finally, the feasibility and effectiveness of the proposed scheme are demonstrated through a case study on the waste water treatment plants (WWTPs).

Keywords: Distributed fault detection, Large-scale systems, Distributed Kalman filter, Weighted average consensus, Expectation-Maximization algorithm.

1. INTRODUCTION

Over the past few years, fault detection (FD) for large-scale interconnected systems, is one of the challenging topics in the field of process monitoring and has gained widespread attention from both academic and industrial application areas (Li et al., 2020; Zhu et al., 2023). Most of the existing FD methods for large-scale systems follow the centralized strategy, where all the local process information are collected in a central unit to perform the corresponding calculation and FD actions (Fanti et al., 2012). Such centralized methods may lead to significant requirements on the computational capacity and data transmitting efforts, and thus result in efficiency and security worries (Ge and Chen., 2015; Yang et al., 2021).

Thanks to the swift progess in computer science and communication technology, current technical systems are often equipped with excellent infrastructure such as sensor networks for data acquisition and management. In light of this, a surge in the research on decentralized FD methods has been stimulated, in order to reduce the communication and computation costs (Grbovic et al., 2012). However, it has been evident that due to neglect of the information exchange among different nodes, the detection performance of decentralized FD is much less than the performance of centralized FD. Under this circumstance, various distributed FD methods have been proposed, where information exchange and data fusion are conducted between adjacent nodes through a communication network (Krishnamachari and Iyengar., 2004). Among the distributed FD methods, distributed Kalman filter based fault detection has gained widespread usage. For example, a distributed extended Kalman filter based FD method is addressed in (Rigatos et al., 2013) to monitor the condition of electric systems. A fault detection approach using fusing unscented Kalman filter is proposed in (Vafamand et al., 2021) to solve the problem of detecting faults in direct current microgrids with nonlinear loads. Nevertheless, most of the existing distributed Kalman filter based fault detection methods are performed under the assumption that measurement matrices and covariance matrices of noise vectors are known. However, precise information of these matrices

⋆ Corresponding author: Xu Yang. (e-mail: yangxu@ustb.edu.cn). This work was supported in part by the National Natural Science Foundation of China under grants 62373012 and 61903026, the Beijing Natural Science Foundation, China under grant 4212040, the Project for Outstanding Youth Team of USTB (Fundamental Research Funds for the Central Universities) under grant FRF-EYIT-23-06.

of the large-scale interconnected system is often difficult to be obtained due to its high order and distributed characteristics. Study on this issue is the major objective of this paper.

Towards information exchange in the distributed fault detection methods, the average consensus algorithm is one of the most widely-used approaches to deal with distributed optimization problems (Xiao and Boyd., 2004), where the value of each node converges to an average value after consensus interation. However, special situations such as node failures, can impact the overall information fusion results. This motivates the second part of our work, in which a weighted average consensus algorithm is proposed.

Considering the above challenges, the contributions, novelties and advantages of this study lie in the following aspects:

(1) The distributed Kalman filter based fault detection scheme, consisting of offline training and online fault detection, is developed by using the average consensus algorithm, in order to reduce the computational and communication load while achieving the identical FD performance of all nodes.
(2) Considering that the measurement matrices and covariance matrices of noise vectors are difficult to be obtained, distributed Kalman filters are designed in parallel at each sensor node based on the Expectation-Maximization algorithm.
(3) A weighted average consensus algorithm is developed during online fault detection, where higher weights are granted to nodes with smaller deviations from the average value, thus significantly reducing the impact of node failures on global information fusion and enhancing the robustness of fault detection.

This paper is organized as follows. In Section 2, the preliminaries and problem formulation are given. Section 3 proposes a distributed Kalman filter based fault detection scheme incorporating weighted average consensus algorithm. The feasibility of the proposed scheme is validated through the WWTPs benchmark in Section 4. Finally, some conclusions and future works are summarized in Section 5.

2. PRELIMINARIES AND PROBLEM STATEMENT

2.1 Graph theory

Consider a large-scale interconnected system with a sensor network, whose topological structure can be represented using an undirected graph. The set of the sensors is denoted by $\mathcal{N}$ and the set of connections is denoted by E, which can also be called the edges. Edge (i, j) means that sensor i and j are connected. Thus, the graph of the sensor network M sensors can be formulated as follows:

$$\begin{aligned} &\mathcal{G} = (\mathcal{N}, E), \mathcal{N} = \{1, \cdots, M\} \\ &E = \{(i, j) \mid i, j \in \mathcal{N}, i \neq j, \text{ they are connected}\} \end{aligned} \quad (1)$$

As a result, the set of neighbors for the i-th sensor, comprising all sensors connected to it, can be expressed by $\mathcal{N}_i$. That is

$$\mathcal{N}_i = \{j \mid \text{sensor } j \text{ is connected to sensor } i, \ j = 1, \cdots, M\} \quad (2)$$

2.2 Distributed average consensus algorithm

Based on the graph theory, the basic average consensus algorithm is briefly introduced. Considering a network with M sensors, the average consensus algorithm is a method for the iteration of vector $x_i \in \mathcal{R}^{1\times m}$ at the i-th sensor as:

$$x_{i,k+1} = w_{ii}x_{i,k} + \sum_{j\in N_i} w_{ij}x_{j,k}, \ i = 1, \cdots M \quad (3)$$

Setting

$$X_k = \begin{bmatrix} x_{1,k} \\ \vdots \\ x_{M,k} \end{bmatrix}, W = \begin{bmatrix} w_{11} & \cdots & w_{1M} \\ \vdots & \ddots & \vdots \\ w_{M1} & \cdots & w_{MM} \end{bmatrix} \quad (4)$$

The iteration of the nodes can be written as

$$X_{k+1} = WX_k \quad (5)$$

It is claimed that the average consensus has been attained when

$$\lim_{k\to\infty} X_k = \lim_{k\to\infty} W^k X_0 = \frac{\mathbf{1}\mathbf{1}^T}{M} X_0 \quad (6)$$

where $X_0 = [x_{1,0}, x_{2,0}, \cdots x_{M,0}]^T$ is the initial value of X_k, and the final result of the iteration is that the node values converge to the average of their initial values. The determination method of the iteration matrix W is referenced in (Xiao et al., 2007).

2.3 Model description and problem formulation

The dynamics of the considered large-scale interconnected system with a sensor network can be described by

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + f(k), x(0) = x_0, \\ y_i(k) &= C_i x(k) + v_i(k) \in \mathcal{R}^{m_i}, i = 1, \cdots, M \end{aligned} \quad (7)$$

where $x(k) \in \mathbb{R}^m$ is the state vector, A, B, C are the system matrices. $f(k) \in \mathbb{R}^m$ denotes the fault vector and $v_i(k) \sim \mathcal{N}(0, \Sigma_{v_i})$ denotes the measurement noise. In the process, M nodes are tasked with monitoring the system state to ensure the fault detection performance.

In practical industrial applications, the sampling rates may vary among different sensor nodes. Differences in sampling frequencies will result in asynchronous information, affecting the efficiency of data fusion. Therefore, the lifting technique is employed to address this issue.

Define the original sample time of the state to be T_s, and the sample time of the i-th sensor can be defined as $\gamma_i T_s(\gamma_i = 1, 2, \cdots)$. Consequently, the lifted state sample time can be denoted by $\xi = \eta T_s$ and the new state transform matrix A_l is supposed to be A^η, where η can be the common multiple of $\gamma_i(i = 1, 2 \cdots M)$. For ease of presentation, it is assumed that the fault vector $f(\xi)$ remains constant in a lifted sample time ξ. That is $f(\xi) = [ff \cdots f]^T$. To this end, the lifted interconnected system can be denoted by

$$\begin{aligned} x(\xi+1) &= A_l x(\xi) + B_{lu}u + B_{lf}f \\ y_{li}(\xi) &= C_{li}x(\xi) + H_{Bi}u + H_{fi}f(\xi) + v_{li}(\xi) \end{aligned} \quad (8)$$

where

$$B_{lf} = [A^{\eta-1} \cdots A \; I] \qquad B_{lu} = [A^{\eta-1}B \cdots AB \; B] \tag{9}$$

$$H_{fi} = \begin{bmatrix} C_i \\ \sum_{j=0}^{\gamma_i} C_i A^j \\ \vdots \\ \sum_{j=0}^{(\eta_i-1)\gamma_i} C_i A^j \end{bmatrix} \tag{10}$$

$$H_{Bi} = \begin{bmatrix} 0 & \cdots & & & & 0 \\ C_i A^{\gamma_i-1}B & \cdots C_iAB \; C_iB & 0 & \cdots & & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ C_i A^{(\eta_i-1)\gamma_i-1}B & \cdots & \cdots & C_iAB & C_iB & 0 \end{bmatrix} \tag{11}$$

$$C_{li} = \begin{bmatrix} C_i \\ C_iA^{\gamma_i} \\ C_iA^{2\gamma_i} \\ \vdots \\ C_iA^{(\eta_i-1)\gamma_i} \end{bmatrix}, v_{li}(\xi) = \begin{bmatrix} v_i(\xi) \\ v_i(\xi+\gamma_i T_s) \\ v_i(\xi+2\gamma_i T_s) \\ \vdots \\ v_i(\xi+(\eta_i-1)\gamma_i T_s) \end{bmatrix} \tag{12}$$

For sensor node $i = 1, 2, \cdots M$, the lifted system can be summarized as:

$$y_l(\xi) = C_l x(\xi) + H_B u + H_f f_l(\xi) + v_l(\xi), C_l = \begin{bmatrix} C_{l1} \\ C_{l2} \\ \vdots \\ C_{lM} \end{bmatrix}$$

$$H_f = \begin{bmatrix} H_{f1} \\ H_{f2} \\ \vdots \\ H_{fM} \end{bmatrix}, v_l(\xi) = \begin{bmatrix} v_{l1}(\xi) \\ v_{l2}(\xi) \\ \vdots \\ v_{lM}(\xi) \end{bmatrix}, H_B = \begin{bmatrix} H_{B1} \\ H_{B2} \\ \vdots \\ H_{BM} \end{bmatrix} \tag{13}$$

Utilizing the lifted technique, the problem of this study can be formulated as follows: develop a distributed fault detection scheme for the lifted interconnected system (13).

3. A DISTRIBUTED KALMAN FILTER-BASED FD SCHEME INTEGRATING WEIGHTED AVERAGE CONSENSUS ALGORITHM

In this section, a distributed KF based fault detection scheme is addressed for the lifted interconnected system (13). During both the offline training and online detection phase, the average consensus algorithm is implemented for data fusion and simplification of calculations.

3.1 Kalman filter based residual generator for fault detection

Kalman filter is considered an optimal state estimator for linear systems, provided that the system and measurement noises follow Gaussian distributions. For system (13), Kalman filter is conducted at each node as follows:

$$\begin{aligned} \hat{x}(\xi+1) &= A_l\hat{x}(\xi) + B_{lu}u(\xi) + K(\xi)r_l(\xi) \\ \hat{y}(\xi) &= C_l\hat{x}(\xi), r_l(\xi) = y_l(\xi) - \hat{y}_l(\xi) \end{aligned} \tag{14}$$

where $K(\xi)$ denotes the Kalman gain, and tends to be constant when the linear system is in a steady state, which satisfies

$$\begin{aligned} K(\xi) &= A_l\Sigma(\xi)C_l^T\Sigma_r^{-1} \\ \Sigma(\xi) &= A_l\Sigma(\xi)A_l^T - K(\xi)\Sigma_r K(\xi)^T \\ \Sigma_r &= \mathcal{E}(r_l(\xi)r_l^T(\xi)) = C_l\Sigma(\xi)C_l^T + \Sigma_{v_l} \\ \Sigma_{v_l} &= diag(\Sigma_{v_{l1}}, , \Sigma_{v_{l2}, \cdots, \Sigma_{v_{lM}}}), \Sigma_{v_{li}} = diag(\Sigma_{v_i}, \cdots, \Sigma_{v_i}) \end{aligned} \tag{15}$$

where $\Sigma(\xi)$ is the covariance matrix of the state estimation error, and Σ_{v_l} denotes the covariance matrix of the noise vector.

The residual vector $r_l(\xi)$ generated by Kalman filter of (14) can be used for fault detection. Due to the whiteness property of the residual vectors, the FD problem of a dynamic process can be approached as fault detection in a statistical process (Ding., 2021). Consequently, the residual vector $r_l(\xi)$ can be written as

$$r_l(\xi) = \begin{cases} \varepsilon_l \sim \mathcal{N}(0, \Sigma_r), \text{fault-free} \\ \varepsilon_l + H_f f_l(\xi), \text{faulty} \end{cases} \tag{16}$$

Similar to the well-known FD methods for a statistical process, the test statistic and threshold can be constructed as follows:

$$\begin{aligned} J &= {r_l}^T(\xi)\left(H_f^T\Sigma_r^{-1}\right)^T\left(H_f^T\Sigma_r^{-1}H_f\right)^{-1}H_f^T\Sigma_r^{-1}r(\xi) \\ J_{th} &= \chi_\alpha^2(k_f), k_f = \eta m \end{aligned} \tag{17}$$

Hence, the logic of fault detection can be represented by:

$$\begin{cases} J - J_{th} > 0 \text{ faulty} \\ J - J_{th} \leq 0 \text{ fault-free} \end{cases} \tag{18}$$

According to (14)-(17), Kalman gain, covariance matrix of the state estimation error and test statistic need to be calculated during the FD process. As previously indicated, centralized strategies demand substantial computational power and communication bandwidth, particularly when dealing with high-dimensional residual vectors and fault state transition matrices. To tackle this challenge, this study focuses on the distributed FD scheme using the average consensus algorithm. In the subsequent sections, a distributed fault detection approach is proposed, which mainly includes: (i) distributed offline training and (ii) distributed online fault detection.

3.2 Distributed offline training

During the distributed offline training, Kalman gain, covariance matrix and test statistic are calculated at each sensor to reduce the computational burden and communication load, where the measurement matrices C_i and covariance matrix of process noise Σ_{v_i} are firstly considered.

Assume that adequate process data $y_{i,k}, k = 1, 2, \cdots M$, has been collected during the fault-free period, and the mean measurement value $\bar{y}_i$ is calculated and saved at each sensor. Consequently, the following calculation is performed.

$$\tilde{y}_{i,k} = y_{i,k} - \bar{y}_i \tag{19}$$

To this end, the expected value of state $\tilde{x} = x - \bar{x}$ can be estimated as

$$E(\tilde{x} \mid \tilde{y}_i) = C_i^T \left(\Sigma_{v_i} + C_i C_i^T\right)^{-1} \tilde{y}_i := P_i \tilde{y}_i \tag{20}$$

where P_i represents the projection coefficient of $\tilde{y}_i$. On this basis, the expected value of $\tilde{x}\tilde{x}^T$ can be written as:

$$E\left(\tilde{x}\tilde{x}^T \mid \tilde{y}_i\right) = I - P_i H_j + P_i y_i y_i^T P_i^T \tag{21}$$

In accordance with (Ghahramani and Hinton., 1997) and (Rubin and Thayer., 1982), the estimation of C_i and Σ_{vi} can be carried out at each sensor through the implementation of the Expectation-Maximization algorithm, as outlined in *Algorithm 1.*

Algorithm 1 Estimation of C_i,Σ_{v_i} with EM algorithm

1. Initialize C_i and Σ_i;
2. Iterate the EM step described below until convergence is achieved:

   E step: Compute (20) and (21) at each node;

   M step: Update C_i and Σ_{v_i} by

$$C_i^{new} = \left(\sum_{k=1}^{N} \tilde{y}_{i,k} E\left(\tilde{x} \mid \tilde{y}_{i,k}\right)^T\right)\left(\sum_{k=1}^{N} E\left(\tilde{x}\tilde{x}^T \mid \tilde{y}_{i,k}\right)\right)^{-1}$$

$$\Sigma_{v_i}^{new} = \frac{1}{N} \operatorname{diag}\left\{\sum_{k=1}^{N}\left(\tilde{y}_{i,k}\tilde{y}_{i,k}^T - C_i^{\text{new}} E\left(\tilde{x} \mid \tilde{y}_{i,k}\right)\tilde{y}_{i,k}^T\right)\right\} \tag{22}$$

As a result, C_{li} in (12) and $\Sigma_{v_{li}}$ in (15) can be calculated at each sensor.

To calculate $\Sigma(\xi)$ in a distributed fashion, rewrite the equation in (15) as

$$\begin{aligned}\Sigma(\xi) &= A_l \Sigma(\xi) A_l^T - A_l \Sigma(\xi) C_l^T (C_l \Sigma(\xi) C_l^T)^{-1} C_l \Sigma(\xi) A_l \\ &= A_l (\Sigma(\xi)^{-1} + \sum_{i=1}^{M} C_{li}^T \Sigma_{vli}^{-1} C_{li})^{-1} A_l^T \\ &= A_l (\Sigma(\xi)^{-1} + \sum_{i=1}^{M} \Psi_i)^{-1} A_l^T,\end{aligned} \tag{23}$$

$$\Psi_i = C_{li}^T \Sigma_{vli}^{-1} C_{li}, \bar{\Psi}_i = \frac{1}{M}\sum_{i=1}^{M}\Psi_i, \sum_{i=1}^{M} C_{li}^T \Sigma_{vli}^{-1} C_{li} = M\bar{\Psi}_i \tag{24}$$

where $\bar{\Psi}_i$ represents the average value of Ψ_i and is calculated through the average consensus algorithm. For the computation of $\bar{\Psi}_i$, the average consensus applied with the interation of Ψ_i at the i-th sensor as

$$\Psi_{i,k+1} = w_{ii}\Psi_{i,k} + \sum_{j\in N_i} w_{ij}\Psi_{j,k},\ i = 1, \cdots M \tag{25}$$

is performed started with $\Psi_{i,0} = \Psi_i$ until convergence. On this basis, the covariance matrix of the state estimation error can be calculated through the Riccati equation (23).

Afterwards, Kalman gain $K(\xi)$ and parameters for the test statistic can be computed in parallel at each node. According to (15), (17), and (23), the calculation of $K(\xi)$ and test statistic require the computations of

$$C_l^T \Sigma_r^{-1}, H_f^T \Sigma_r^{-1} H_f, H_f^T \Sigma_r^{-1} \tag{26}$$

Based on (15), it is obvious that

$$\begin{aligned}\Sigma_r^{-1} &= (C_l \Sigma(\xi) C_l^T + \Sigma_{v_l})^{-1} \\ &= \Sigma_{v_l}^{-1} - \Sigma_{v_l}^{-1} C_l (\Sigma(\xi)^{-1} + C_l^T \Sigma_{v_l}^{-1} C_l)^{-1} C_l^T \Sigma_{v_l}^{-1} \\ &= \Sigma_{v_l}^{-1} - \Sigma_{v_l}^{-1} C_l (\Sigma(\xi)^{-1} + M\bar{\Psi}_i)^{-1} C_l^T \Sigma_{v_l}^{-1}\end{aligned} \tag{27}$$

$$C_l^T \Sigma_{v_l}^{-1} = [C_{l1}^T \Sigma_{v_{l1}}^{-1} \cdots C_{lM}^T \Sigma_{v_{lM}}^{-1}] \tag{28}$$

Combined with (27), the parameters of (26) can be calculated as

$$\begin{aligned}H_f^T \Sigma_r^{-1} &= H_f^T \Sigma_{vl}^{-1} - H_f^T \Sigma_{vl}^{-1} C_l \\ &\quad \cdot (\Sigma(\xi)^{-1} + M\bar{\Psi}_i)^{-1} C_l^T \Sigma_{vl}^{-1} \\ H_f^T \Sigma_{vl}^{-1} &= [H_{l1}^T \Sigma_{v_{l1}}^{-1} \cdots H_{lM}^T \Sigma_{v_{lM}}^{-1}] \\ H_f^T \Sigma_{vl}^{-1} C_l &= \sum_{i=1}^{M} H_{fi}^T \Sigma_{vli}^{-1} C_{li} = \sum_{i=1}^{M} \Theta_i\end{aligned} \tag{29}$$

$$H_{fi}^T \Sigma_{vli}^{-1} C_{li} = \Theta_i, \bar{\Theta}_i = \frac{1}{M}\sum_{i=1}^{M}\Theta_i, \sum_{i=1}^{M} H_{fi}^T \Sigma_{vli}^{-1} C_{li} = M\bar{\Theta}_i \tag{30}$$

$$\begin{aligned}H_f^T \Sigma_r^{-1} H_f &= H_f^T \Sigma_{vl}^{-1} H_f - M\bar{\Theta}_i \\ &\quad \cdot (\Sigma(\xi)^{-1} + M\bar{\Psi}_i)^{-1} (M\bar{\Theta}_i)^T \\ H_f^T \Sigma_{vl}^{-1} H_f &= \sum_{i=1}^{M} H_{fi}^T \Sigma_{vli}^{-1} H_{fi} = \sum_{i=1}^{M} \Upsilon_i\end{aligned} \tag{31}$$

$$H_{fi}^T \Sigma_{vli}^{-1} H_{fi} = \Upsilon_i, \bar{\Upsilon}_i = \frac{1}{M}\sum_{i=1}^{M}\Upsilon_i, \sum_{i=1}^{M} H_{fi}^T \Sigma_{vli}^{-1} H_{fi} = M\bar{\Upsilon}_i \tag{32}$$

where $\bar{\Theta}_i$ and $\bar{\Upsilon}_i$ are the average values of Θ_i and Υ_i respectively and can be calculated with the average consensus algorithm like (25). In summary, *Algorithm 2* outlines the proposed offline training phase.

Algorithm 2 Distributed offline training

1. Compute the following parameters at each node:
$$C_{li}^T \Sigma_{vli}^{-1} C_{li},\quad H_{fi}^T \Sigma_{vli}^{-1} C_{li},\quad H_{fi}^T \Sigma_{vli}^{-1} H_{fi} \tag{33}$$
2. Derive the average values of (33) through the utilization of average consensus algorithm.
3. Resolve the Riccati equation as stipulated in (23) in node $1, 2 \ldots M$.
4. Calculate the parameters of (26) to obtain Kalman gain and test statistic in parallel at each node.

3.3 Distributed online fault detection

In the online detection phase, the residuals $r_j(\xi)$ are generated in parallel at all the nodes, and the distributed realisation of residual generator in (14) is written as

$$\begin{aligned}\hat{x}(\xi+1) &= A_l \hat{x}(\xi) + B_{lu} u(k) + K(\xi) r(k) \\ &= A_l \hat{x}(\xi) + B_{lu} u(k) + A_l \Phi^{-1} \sum_{i=1}^{M} C_{li}^T \Sigma_{vl}^{-1} r_i(\xi) \\ \Phi &= \Sigma(k)^{-1} + C_l^T \Sigma_{vl}^{-1} C_l = \Sigma(k)^{-1} + M\bar{\Psi}_i\end{aligned} \tag{34}$$

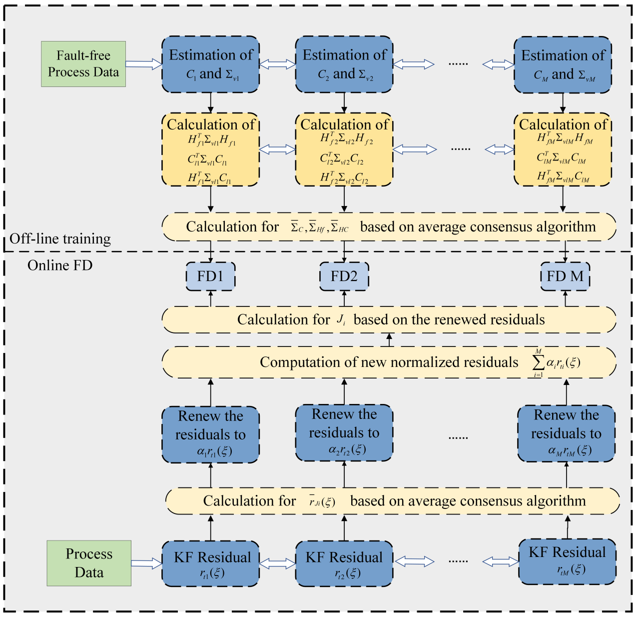

Fig. 1. The schematic of distributed FD with weighted average consensus algorithm

$$r_i(\xi) = y_{li}(\xi) - \hat{y}_{li}(\xi), \hat{y}_{li}(\xi) = C_{li}\hat{x}(\xi)$$
$$r_l(\xi) = [r_1(\xi), r_2(\xi), \cdots, r_M(\xi)]^T \tag{35}$$

Combined with (34), the test statistic at each node can be written as:

$$J_i = {r_l}^T(\xi)\left(H_f^T\Sigma_r^{-1}\right)^T\left(H_f^T\Sigma_r^{-1}H_f\right)^{-1}H_f^T\Sigma_r^{-1}r_l(\xi)$$
$$= r_t^T(\xi)\left(H_f^T\Sigma_r^{-1}H_f\right)^{-1}r_t(\xi)$$
$$r_t(\xi) = H_f^T\Sigma_r^{-1}r_l(\xi)$$
$$= \sum_{i=1}^{M}(H_{fi}^T - M\bar{\Theta}_i\Phi^{-1}C_{li}^T)\Sigma_{vli}^{-1}r_i(\xi) = \sum_{i=1}^{M} r_{ti}(\xi)$$
$$r_{ti}(\xi) = (H_{fi}^T - M\bar{\Theta}_i\Phi^{-1}C_{li}^T)\Sigma_{vli}^{-1}r_i(\xi)$$
$$\bar{r}_{ti}(\xi) = \frac{1}{M}\sum_{i=1}^{M} r_{ti}(\xi), r_t(\xi) = M\bar{r}_{ti}(\xi) \tag{36}$$

Also, the value of $r_t(\xi)$ can be calculated using average consensus algorithm:

$$r_{ti,k+1} = w_{ii}r_{ti,k} + \sum_{j\in N_i} w_{ij}r_{tj,k}, \ i = 1, \cdots M \tag{37}$$

The procedure for online distributed Kalman filter-based fault detection method is summarized in *Algorithm 3*.

Algorithm 3 Distributed online fault detection

1. Construct the residual signals $r_i(\xi)$ based on the distributed Kalman filters.
2. Utilize the average consensus algorithm to compute J_i at each node.
3. Run the fault detection logic at all nodes according to (18).

3.4 The weighted average consensus algorithm

The above distributed fault method is developed on the assumption that all the sensors operate under normal conditions. However, in the event of sensor communication being normal but some sensors losing their sensing capabilities, erroneous information may be fused.

In addressing this situation, each sensor will store its initial normalized residual $r_{ti}(\xi)$ and compare it with the average value $\bar{r}_{ti}(\xi)$. Each node is assigned a weight α_i based on its deviation from the average value, so that the measurement results of nodes with larger deviations from the average have less impact on the overall calculation result. That is:

$$r_{ti}(\xi)_{new} = \alpha_i r_{ti}(\xi)$$
$$\alpha_i = \frac{1}{1 + |r_{ti}(\xi) - \bar{r}_{ti}(\xi)|} \Big/ \sum_{i=1}^{M}\frac{1}{1 + |r_{ti}(\xi) - \bar{r}_{ti}(\xi)|} \tag{38}$$

The new normalized residuals at each node can be rewritten as $\bar{r}_{ti}(\xi) = \sum_{i=1}^{M}\alpha_i r_{ti}(\xi)$. In summary, the schematic of the proposed distributed fault detection scheme incorporating a weighted average consensus algorithm is shown in Fig. 1

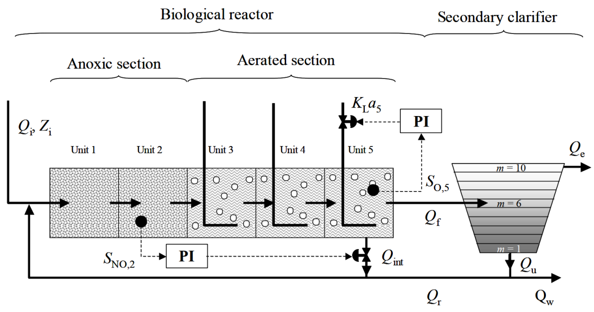

Fig. 2. The reaction process of the WWTP

4. IMPLEMENTATION RESULT

The effectiveness and performance of the proposed distributed fault detection scheme are now illustrated by a case study on the the waste water treatment plants (WWTPs), whose reaction process is depicted in Fig. 2. A senor network is equipped in the 5th Bioreactor with the graph as (39) to measure the the dissolved oxygen $S_{O,5}$.

$$\Pi = [(1,2), (2,3), (3,4), (3,5), (4,5), (1,4)] \tag{39}$$

Therefore, According to (Xiao et al., 2007), the iteration matrix W of average consensus algorithm can be formulated as

$$W = \begin{bmatrix} 0.5143 & 0.2428 & 0 & 0.2428 & 0 \\ 0.2428 & 0.5143 & 0.2428 & 0 & 0 \\ 0 & 0.2428 & 0.2715 & 0.2428 & 0.2428 \\ 0.2428 & 0 & 0.2428 & 0.2715 & 0.2428 \\ 0 & 0 & 0.2428 & 0.2428 & 0.5143 \end{bmatrix} \tag{40}$$

For demonstration purpose, two fault scenarios are performed in this study.

In case I, a stepwise fault is inserted from the 400th timestep, leading to an increase in the $S_{O,5}$ within the 5th Bioreactor. Fig. 3 shows the results of centralized, decentralized and distributed Kalman filter based FD methods. It is clear that the decentralized approach exhibits a relatively higher missed detection rate (MDR), while the distributed approach demonstrates similar fault

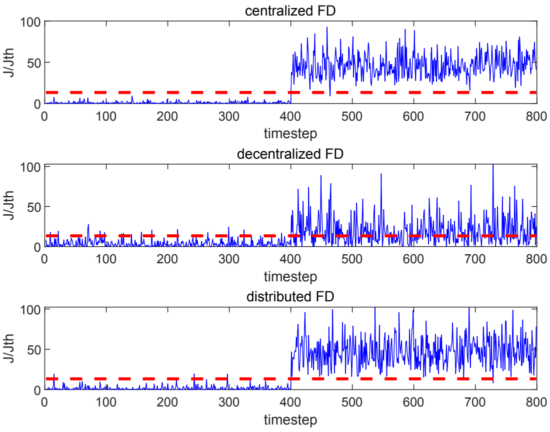

Fig. 3. Comparison of centralized, decentralized, and distributed fault detection algorithms

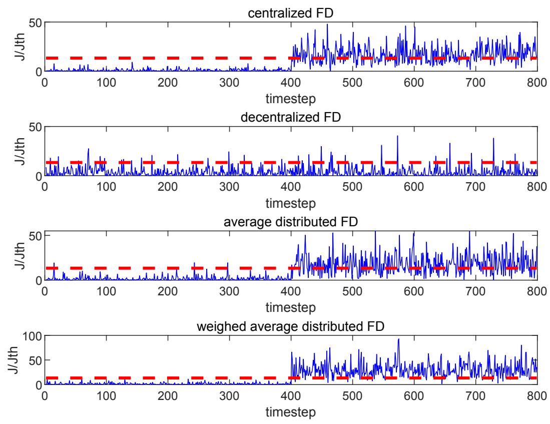

Fig. 4. Comparison of fault detection results before and after failure of node 1

detection performance as the centralized method. Meanwhile, the distributed FD method significantly reduces computational and communication load since it splitts the centralized high-dimensional matrices for computation into lower-dimensional matrices.

In case II, failure of node 1 is generated at the 400th timestep. The second sub-figure in Fig. 4 illustrates the MDR of the decentralized fault detection method has reached a significant level due to the failure of node 1. Meanwhile, the fault detection results for the centralized FD method and the distributed FD method are depicted in the first and third sub-figures in Fig. 4, MDR of which are 35.75% and 37% respectively. After incorporating the weighted average consensus algorithm, as shown in the last sub-figure in Fig. 4, the MDR has significantly decreased, reaching as low as 10.5%.

5. CONCLUSION

In this paper, a distributed Kalman filter based fault detection scheme for large-scale systems with a sensor network have been proposed. To this end, the Expectation-Maximization algorithm has been introduced to identify the measurement matrices and covariance matrices of process noise. On this basis, the distributed Kalman filter based residual generator and test statistic have been designed using the average consensus algorithm, reducing the computational load and enhancing the reliability of fault detection. It is followed by a weighted average consensus based distributed fault detection scheme to deal with some special conditions, such as the occurrence of node failures. Our future work will dedicate to distributed fault detection for nonlinear large-scale systems.

REFERENCES

Ding, S. X. (2021). Advanced methods for fault diagnosis and fault-tolerant control. *Springer*.

Fanti, M. P., Mangini, A. M., and Ukovich, W. (2012). Fault detection by labeled Petri nets in centralized and distributed approaches. *IEEE Transactions on Automation Science and Engineering*, 10(2), 125–247.

Ge, Z., and Chen, J. (2015). Plant-wide industrial process monitoring: A distributed modeling framework. *IEEE Transactions on Industrial Informatics*, 12(1), 2015.

Ghahramani, Z., and Hinton, G. E. (1997). The EM algorithm for mixtures of factor analyzers. *Technical Report CRG-TR-96-1*, University of Toronto.

Grbovic, M., Li, W., Xu, P., Usadi, A. K., Song, L., and Vucetic, S. (2012). Decentralized fault detection and diagnosis via sparse PCA based decomposition and maximum entropy decision fusion. *Journal of Process Control*, 22(4), 738-750.

Krishnamachari, B., and Iyengar, S. (2004). Distributed Bayesian algorithms for fault-tolerant event region detection in wireless sensor networks. *IEEE Transactions on Computers*, 53(3), 241-250.

Li, L., Ding, S. X., and Peng, X. (2020). Distributed data-driven optimal fault detection for large-scale systems. *Journal of Process Control*, 96, 94–103.

Rigatos, G., Siano, P., and Zervos, N. (2013). A distributed state estimation approach to condition monitoring of nonlinear electric power systems. *Asian Journal of Control*, 15(3), 849-860.

Rubin, D. B., and Thayer, D. T. (1982). EM algorithms for ML factor analysis. *Psychometrika*, 47, 69-76.

Vafamand, A., Moshiri, B., and Vafamand, N. (2021). Fusing unscented Kalman filter to detect and isolate sensor faults in DC microgrids with CPLs. *IEEE Transactions on Instrumentation and Measurement*, 71, 1-8.

Xiao, L., and Boyd, S. (2004). Fast linear iterations for distributed averaging. *Systems & Control Letters*, 53(1), 65–78.

Xiao, L., Boyd, S., and Kim, S. J. (2007). Distributed average consensus with least-mean-square deviation. *Journal of parallel and distributed computing*, 67(1), 33-46.

Yang, X., Gao, J., and Huang, B. (2021). Data-driven design of fault detection and isolation method for distributed homogeneous systems. *Journal of the Franklin Institute*, 358(9), 4929-4949.

Zhu, M., Sui, T., and Wang, R. (2023). Distributed Kalman filtering over sensor networks with fading measurements and random link failures. *Journal of the Franklin Institute*, 360, 2457–2475.