Fault Detection and Isolation of Uncertain Nonlinear Parabolic PDE Systems

Jingting Zhang, Chengzhi Yuan, Wei Zeng, Cong Wang

***Abstract*—This paper proposes a novel fault detection and isolation (FDI) scheme for distributed parameter systems modeled by a class of parabolic partial differential equations (PDEs) with nonlinear uncertain dynamics. A key feature of the proposed FDI scheme is its capability of dealing with the effects of system uncertainties for accurate FDI. Specifically, an approximate ordinary differential equation (ODE) system is first derived to capture the dominant dynamics of the original PDE system. An adaptive dynamics identification approach using radial basis function neural network is then proposed based on this ODE system, so as to achieve locally-accurate identification of the uncertain system dynamics under normal and faulty modes. A bank of FDI estimators with associated adaptive thresholds are finally designed for real-time FDI decision making. Rigorous analysis on the FDI performance in terms of fault detectability and isolatability is provided. Simulation study on a representative transport-reaction process is conducted to demonstrate the effectiveness and advantage of the proposed approach.**

***Index Terms*—Partial differential equations, fault detection and isolation, adaptive dynamics identification, deterministic learning, neural network, distributed parameter systems.**

I. Introduction

DISTRIBUTED parameter systems (DPSs) are dynamical systems with inputs, outputs, and process parameters that may vary temporally and spatially [1], [2], [3], which are usually modeled by partial differential equations (PDEs). Some typical examples include fluid flow process [4], biological process [5], convection diffusion reaction process [6] and thermal process [7]. Particularly, due to the ever-increasing technical demands, fault diagnosis of DPSs has been an area of significantly growing interests. It is a critical step to realize fault tolerant operations for minimizing performance degradation and avoiding dangerous situations, such that safety and reliability of DPSs can be guaranteed. To this end, the past decades have witnessed tremendous progress in the research of fault diagnosis for DPSs, leading to a large variety of methods, see, e.g., [8], [9], [10], [11], [12] and the references therein.

As opposed to the substantially growing body of literature on fault detection (FD) of DPSs (e.g., [8], [9], [10], [13], [14]), study on the fault isolation (FI) problem has gained quite limited success, especially for those DPSs with nonlinear unstructured uncertain dynamics. Some research efforts have been devoted to the development of FI methods for DPSs with precisely known models. For example, [15] proposed an FI scheme using a finite-dimensional geometric approach. In [6], the FI problem for DPSs with various actuator faults has been investigated. For the FI problem of DPSs with nonlinear uncertain dynamics, the research is still under-explored. One of the technical difficulties is that the dynamics of faults occurring in the system could be hidden within the system's general uncertain dynamics (e.g., unmodeled uncertainties), such that the fault feature could not be accurately identified for FI purpose. Some attempts have been made to overcome this difficulty. The FI scheme proposed in [16] is able to distinguish the effects between occurring fault and system uncertainties. [17], [18] developed a Lyapunov function-based FI scheme for DPSs, in which system uncertainties were handled by active control strategies. However, all these existing schemes have not appropriately dealt with the system uncertainties in the sense that occurring faults are typically required to have sufficiently large magnitudes (e.g., larger than those of the system uncertainties), limiting their wider applicability in practice.

To overcome the above deficiencies, a promising strategy is to realize accurate modeling of the system uncertain dynamics. Adaptive neural network (NN)–as commonly used in the field of control and modeling of DPSs with uncertain dynamics (see, e.g., [19], [20], [21])–provides a powerful tool for this purpose. However, different from the control problem, where modeling errors of NN can typically be compensated by the controllers, using NN for accurate fault detection and isolation (FDI) of DPSs is rather challenging. This is because NN approximation errors often have negative impacts on the FDI residual signals, which cannot be structured and decoupled from the occurring fault, leading to possible misjudgment of FDI. To minimize such effects of NN approximation errors, a key technical challenge is to satisfy the so-called persistently exciting (PE) condition of associated NN regressor vectors [22]. Recently, the deterministic learning (DL) theory proposed in [23] has demonstrated that with radial basis function neural networks (RBF NN), almost any recurrent trajectory can result in the satisfaction of a partial PE condition. As a result, locally accurate RBF NN identification of nonlinear unstructured uncertain dynamics can be achieved along the recurrent trajectory, and approximation errors of the associated NN can be guaranteed arbitrarily small [24]. With this important property, the DL theory has been recognized in recent years as a new and effective paradigm for the design of FDI schemes for general nonlinear uncertain systems, see

This work was supported in part by the National Science Foundation under Grant CMMI-1929729.

J. Zhang and C. Yuan are both with the Department of Mechanical, Industrial and Systems Engineering, University of Rhode Island, Kingston, RI 02881, USA (e-mail: jingting_zhang@uri.edu; cyuan@uri.edu)

W. Zeng is with the School of Mechanical and Eletrical Engineering, Longyan University, Longyan 364012, China (e-mail: zw0597@126.com)

C. Wang is with the School of Control Science and Engineering, Shandong University, Jinan 250061, China (e-mail: wangcong@sdu.edu.cn)

[25], [26], [27], [28]. This has opened new doors to the field though, existing DL-based FDI schemes still suffer from some limitations. For example, the FI methods of [25], [26], [27] require the dynamics of occurring faults to perfectly match those pre-defined/pre-trained faults. However, in many practical situations, e.g., when the system suffers disturbances resulting in changes of initial conditions or system parameters, the dynamics of occurring fault often exhibit some differences from those of the matched fault, which thus could result in missed/false alarm under the FI schemes of [25], [26], [27]. Furthermore, virtually all of these existing FDI methods are focused on lumped parameter systems (LPSs) modeled by finite-dimensional ordinary differential equations (ODEs), which cannot be directly applied for infinite-dimensional DPSs as considered in the current paper.

To extend the DL theory to the FDI problem of DPSs, one promising strategy is to use a finite set of ODEs to approximate the PDE model based on model reduction methods [29]. Conventional spatial discretization-based approaches often lead to a high-order ODE system [30], which however could be computationally expensive for real-time implementation. An alternative approach is based on the Galerkin method [2], [19], [31]. The key idea of the Galerkin method is as follows. It is known that for dissipative parabolic DPSs, the eigenspectrum of the associated spatial differential operator can be partitioned into a finite set of slow eigenvalues and an infinite set of fast but stable eigenvalues [34]. By neglecting the fast stable components, a low-order ODE system can be obtained to approximate the dominant dynamics of the PDE system, which then could be utilized to facilitate the subsequent FDI design. It is worth mentioning that for FDI of DPSs, only a few research results have been obtained in [15], [6], [17], [18], which unfortunately are applicable only to some special types of faults such as actuator faults, but cannot address the FDI problem for DPSs with more general faults.

In this paper, we aim to investigate effective detection and isolation approaches for general faults occurring in DPSs modeled by a class of parabolic PDEs with nonlinear unstructured uncertain dynamics. A novel Galerkin-DL-based FDI scheme will be proposed. Specifically, with the Galerkin method, an approximate ODE model is first derived to capture the dominant dynamics of the PDE system. A DL-based adaptive dynamics identification approach is then developed based on this ODE system to realize locally-accurate identification of the uncertain system dynamics under normal and all faulty modes. The associated knowledge can be obtained and stored in constant RBF NN models. Afterwards, a bank of FDI estimators are designed with these constant models, where the FD estimators are used to detect the occurrence of a fault, and the FI estimators (which will be activated once the occurring fault is detected) are used to identify the type of the occurring fault. Their generated residuals can be used to characterize the dynamics of the occurring fault and distinguish it from the system uncertain dynamics for accurate FDI. Adaptive thresholds associated with such FDI residuals are further designed to facilitate real-time FDI decision making. In particular, to address the aforementioned robustness issue encountered by the FI methods of [25], [26], [27], novel adaptive thresholds instead of fixed/constant thresholds are designed in this paper, such that successful isolation can still be guaranteed even when the occurring fault does not exactly match any of the pre-trained faults. As such, the proposed FI scheme possesses improved robustness against slight deviations of fault dynamics due possibly to unexpected system changes in, for example, initial conditions and system parameters, as discussed above. We stress that our FDI scheme does not require the faults to be of any special type (e.g., actuator faults as required in [17], [18], [15] and/or sensor faults as required in [16]), but is applicable to general system faults. Rigorous analysis on the FDI performance is conducted to demonstrate that our approach develops better fault detectability and isolatability compared to the existing methods of [16], [6], [32], [9]. Moreover, extensive simulations applied to a representative transport-reaction process are also conducted to demonstrate the effectiveness and advantage of the proposed new methodologies.

It should be pointed out that this research work significantly expands our previous work [14] which was focused on only the FD problem of uncertain parabolic PDE systems; while in this paper, we consider both the FD and FI problems. In addition, new adaptive thresholds are proposed for more accurate and efficient FDI, which advance the fixed/constant thresholds proposed in [14]. Furthermore, this paper also provides rigorous analysis to characterize the properties of the proposed FDI scheme, which include: (i) fault detectability conditions characterizing the class of faults that can be detected, and (ii) fault isolatability conditions characterizing the class of faults that can be isolated.

The contributions of this paper are summarized as follows: (i) The FDI problem for uncertain nonlinear parabolic PDE systems is addressed using a novel Galerkin-DL-based adaptive dynamics identification approach, which can achieve locally-accurate identification of the dominant uncertain dynamics of the PDE system; (ii) New adaptive-threshold-based FDI decision making schemes are proposed, which are capable of dealing with general faults occurring in parabolic PDE systems, including those faults that do not exactly match the pre-defined/pre-trained faults; (iii) Rigorous analysis on FDI performance, including fault detectability and isolatability conditions, is provided to demonstrate the effectiveness of the proposed approaches.

The remainder of this paper is organized as follows. Some preliminaries and the problem statement are provided in Section II. The DL-based adaptive dynamics learning approach is presented in Section III. The FD scheme is proposed in Section IV, and the FI scheme is given in Section V. Simulation studies are presented in Section VI. Section VII concludes the paper.

Notation. $\mathbb{R}$, $\mathbb{R}_+$ and $\mathbb{N}_+$ denote, respectively, the set of real numbers, the set of positive real numbers and the set of positive integers; $\mathbb{R}^{m\times n}$ denotes the set of $m \times n$ real matrices; $\mathbb{R}^n$ denotes the set of $n\times 1$ real column vectors; $|\cdot|$ is the absolute value of a real number; $\|\cdot\|$ is the 2-norm of a vector or a matrix, i.e. $\|x\| = (x^\top x)^{\frac{1}{2}}$.

II. Preliminaries and Problem Formulation

A. Radial Basis Function Neural Networks

The RBF NNs can be described by $f_{nn}(Z) = \sum_{i=1}^{N_n} w_i s_i(Z) = W^\top S(Z)$ [33], where $Z \in \Omega_Z \subset \mathbb{R}^q$ is the input vector, $W = [w_1, \cdots, w_{N_n}]^\top \in \mathbb{R}^{N_n}$ is the weight vector, N_n is the NN node number, and $S(Z) = [s_1(\|Z - \varsigma_1\|), \cdots, s_{N_n}(\|Z - \varsigma_{N_n}\|)]^\top$, with $s_i(\cdot)$ being a radial basis function, and ς_i $(i = 1, 2, \cdots, N_n)$ being distinct points in state space. The Gaussian function $s_i(\|Z - \varsigma_i\|) = \exp[\frac{-(Z-\varsigma_i)^\top (Z-\varsigma_i)}{\nu_i^2}]$ is one of the most commonly used radial basis functions, where $\varsigma_i = [\varsigma_{i1}, \varsigma_{i2}, \cdots, \varsigma_{iq}]^\top$ is the center of the receptive field and ν_i is the width of the receptive field. The Gaussian function belongs to the class of localized RBFs in the sense that $s_i(\|Z-\varsigma_i\|) \to 0$ as $\|Z\| \to \infty$. It is noted that $S(Z)$ is bounded, i.e., there exists a real constant $S_M \in \mathbb{R}_+$ such that $\|S(Z)\| \leqslant S_M$ [23, Lemma 2.1]. It has been shown in [33] that for any continuous function $f(Z) : \Omega_Z \to \mathbb{R}$ where $\Omega_Z \subset \mathbb{R}^q$ is a compact set, and for the NN approximator, where the node number N_n is sufficiently large, there exists an ideal constant weight vector W^*, such that for any $\epsilon^* > 0$, $f(Z) = W^{*\top} S(Z) + \epsilon, \forall Z \in \Omega_Z$, where $|\epsilon| < \epsilon^*$ is the ideal approximation error. The ideal weight vector W^* is an "artificial" quantity used for analysis, which is defined as the value of W that minimizes $|\epsilon|$ for all $Z \in \Omega_Z \subset \mathbb{R}^q$, i.e., $W^* \triangleq arg\min_{W \in \mathbb{R}^{N_n}} \{\sup_{Z \in \Omega_Z} |f(Z) - W^\top S(Z)|\}$. Moreover, based on the localization property of RBF NNs [23], for any bounded trajectory $Z(t)$ within the compact set Ω_Z, $f(Z)$ can be approximated by using a finite number of neurons located in a local region along the trajectory: $f(Z) = W_\zeta^{*\top} S_\zeta(Z) + \epsilon_\zeta$, where ϵ_ζ is the approximation error, with $\epsilon_\zeta = O(\epsilon) = O(\epsilon^*)$, $S_\zeta(Z) = [s_{j1}(Z), \cdots, s_{j\zeta}(Z)]^\top \in \mathbb{R}^{N_\zeta}$, $W_\zeta^* = [w_{j1}^*, \cdots, w_{j\zeta}^*]^\top \in \mathbb{R}^{N_\zeta}$, $N_\zeta < N_n$, and the integers $j_i = j_1, \cdots, j_\zeta$ are defined such that $|s_{j_i}(Z_p)| > \theta$ ($\theta > 0$ is a small positive constant) for some $Z_p \in Z(t)$. In addition, it is shown in [23] that for a localized RBF network $W^\top S(Z)$ whose centers are placed on a regular lattice, almost any recurrent trajectory[1] $Z(t)$ can lead to the satisfaction of the PE condition of the regressor subvector $S_\zeta(Z)$. This result is summarized in the following lemma.

Lemma 1 ([23]). *Consider any recurrent trajectory $Z(t)$ that remains in a bounded compact set $\Omega_Z \subset \mathbb{R}^q$. For RBF network $W^\top S(Z)$ with centers placed on a regular lattice (large enough to cover compact set Ω_Z), the regressor subvector $S_\zeta(Z)$ consisting of RBFs with centers located in a small neighborhood of $Z(t)$ satisfies the PE condition.*

B. Problem Formulation

Consider a class of nonlinear parabolic PDE systems in one spatial dimension with a state-space description in the form of:

$$\begin{aligned}\frac{\partial x(z,t)}{\partial t} =& a_1 \frac{\partial x(z,t)}{\partial z} + a_2 \frac{\partial^2 x(z,t)}{\partial z^2} + f(x,u) \\ &+ \beta(t - t_0)\phi^k(x,u),\end{aligned} \tag{1}$$

subject to the following boundary conditions and initial condition:

$$\begin{aligned} &m_i x(z_i,t) + n_i \frac{\partial x}{\partial z}(z_i,t) = d_i, \quad i = 1,2 \\ &x(z,0) = x_0(z),\end{aligned} \tag{2}$$

where $x(z,t) \in \mathbb{R}$ is system state; $u \in \mathbb{R}^q$ is system input; $z \in [z_1, z_2]$ is the spatial coordinate; $t \in [0, \infty)$ is the time; $f(x,u) \in \mathbb{R}$ and $\phi^k(x,u) \in \mathbb{R}$ are unknown nonlinear functions satisfying locally Lipschitz continuous, which represent nonlinear uncertain system dynamics and deviations in system dynamics due to fault $k \in \{1, 2, \cdots, N\}$ ($N \in \mathbb{N}_+$), respectively; $\beta(t - t_0)$ is the time profile of the occurring fault, with $\beta(t-t_0) = 0$ for $t < t_0$ and $\beta(t-t_0) = 1$ for $t \geq t_0$; t_0 is the unknown fault occurrence instant; $\frac{\partial x}{\partial z}$ and $\frac{\partial^2 x}{\partial z^2}$ are the first-order and second-order spatial derivatives of $x(z,t)$, respectively; $a_1, a_2, m_1, m_2, n_1, n_2, d_1, d_2$ are known constants. In this paper, it is assumed that the system state $x(z,t)$ is measurable at all locations $z \in [z_1, z_2]$ for all time $t \in [0, \infty)$.

Assumption 1. *For the PDE system (1)–(2), the system input $u(t)$ and state $x(z_0,t)$ at any spatial point $z_0 \in [z_1, z_2]$ are bounded and recurrent for all $t \in [0, \infty)$.*

Denote $\mathcal{H}$ as a Hilbert space of 1-D functions defined on $[z_1, z_2]$ that satisfies the boundary conditions (2), with inner product and norm: $\langle \zeta_1, \zeta_2 \rangle = \int_{z_1}^{z_2} \zeta_1(z)\zeta_2(z)dz$, $\|\zeta_1\|_2 = \langle \zeta_1, \zeta_1 \rangle^{\frac{1}{2}}$, where $\zeta_1(z)$, $\zeta_2(z)$ are two elements of $\mathcal{H}$. According to [6], [19], the PDE system (1)–(2) can be formulated as an infinite-dimensional system:

$$\dot{\chi} = \mathcal{A}\chi + f(\chi, u) + \beta(t - t_0)\phi^k(\chi, u), \quad \chi(0) = \chi_0, \tag{3}$$

where $\chi(t) = x(z,t)$ is the state function defined in $\mathcal{H}$, and $\mathcal{A}$ is a differential operator in $\mathcal{H}$ defined as $\mathcal{A}x = a_1 \frac{\partial x}{\partial z} + a_2 \frac{\partial^2 x}{\partial z^2}$, $x \in D(\mathcal{A}) := \{x \in \mathcal{H} \,|\, \mathcal{A}x \in \mathcal{H}, m_i x(z_i,t) + n_i \frac{\partial x}{\partial z}(z_i,t) = d_i, i = 1,2\}$. For the operator $\mathcal{A}$, the eigenvalue problem is defined as $\mathcal{A}\varphi_j = \lambda_j \varphi_j$ $(j = 1, 2, \cdots, \infty)$, where λ_j denotes an eigenvalue, and φ_j denotes an eigenfunction. The eigenspectrum of $\mathcal{A}$, denoted by $\sigma(\mathcal{A})$, is defined as the set of all eigenvalues of $\mathcal{A}$, i.e., $\sigma(\mathcal{A}) = \{\lambda_1, \lambda_2, \cdots, \lambda_\infty\}$. According to [6], [19], for highly-dissipative PDE systems, the eigenspectrum of $\mathcal{A}$ can be partitioned into a finite-dimensional part consisting of m ($m \in \mathbb{N}_+$) slow eigenvalues and a stable infinite-dimensional complement containing the remaining fast eigenvalues, and the separation between the slow and fast eigenvalues of $\mathcal{A}$ is large. These properties can be satisfied by the majority of diffusion-convection-reaction processes [6], and are formalized in the following assumption.

Assumption 2. *(i) $\mathrm{Re}\{\lambda_1\} \geq \mathrm{Re}\{\lambda_2\} \geq \cdots \geq \mathrm{Re}\{\lambda_j\} \geq \cdots$, where $\mathrm{Re}\{\lambda_j\}$ denotes the real part of λ_j; (ii) $\sigma(\mathcal{A})$ can be partitioned as $\sigma(\mathcal{A}) = \sigma_s(\mathcal{A}) + \sigma_f(\mathcal{A})$, where $\sigma_s(\mathcal{A})$ consists of the first m number of eigenvalues, that is, $\sigma_s(\mathcal{A}) = \{\lambda_1, \lambda_2, \cdots, \lambda_m\}$, and $\left|\frac{\mathrm{Re}\{\lambda_1\}}{\mathrm{Re}\{\lambda_m\}}\right| = O(1)$; (iii) $\mathrm{Re}\{\lambda_{m+1}\} < 0$ and $\left|\frac{\mathrm{Re}\{\lambda_m\}}{\mathrm{Re}\{\lambda_{m+1}\}}\right| = O(\iota)$, where $\iota := \left|\frac{\mathrm{Re}\{\lambda_1\}}{\mathrm{Re}\{\lambda_{m+1}\}}\right| < 1$ is a small positive constant.*

[1] A recurrent trajectory represents a large set of periodic and periodic-like trajectories generated from linear/nonlinear dynamics systems. A detailed characterization of recurrent trajectories can be found in [23].

Based on this assumption, consider the decomposition $\mathcal{H} = \mathcal{H}_s \oplus \mathcal{H}_f$, in which $\mathcal{H}_s = \text{span}\{\varphi_1, \varphi_2, \cdots, \varphi_m\}$ denotes the finite dimensional space spanned by the slow eigenfunctions corresponding to $\sigma_s(\mathcal{A})$, and $\mathcal{H}_f = \text{span}\{\varphi_{m+1}, \varphi_{m+2}, \cdots, \varphi_\infty\}$ denotes the infinite dimensional complement one spanned by the fast eigenfunctions corresponding to $\sigma_f(\mathcal{A})$. Under such a decomposition and through separation of time and spatial variables [6], [19], the PDE system (3) can be rewritten in the following equivalent form:

$$\begin{aligned} \dot{x}_s &= A_s x_s + f_s(x_s, \chi_f, u) + \beta(t-t_0)\phi_s^k(x_s, \chi_f, u), \\ \dot{\chi}_f &= A_f \chi_f + f_f(x_s, \chi_f, u) + \beta(t-t_0)\phi_f^k(x_s, \chi_f, u), \\ x_s(0) &= x_{s_0},\ \chi_f(0) = \chi_{f_0}, \end{aligned} \tag{4}$$

where $x_s = [x_{s_1}, \chi_{s_2}, \cdots, x_{s_m}]^\top \in \mathbb{R}^m$, $\chi_f = [\chi_{f_{m+1}}, \cdots, \chi_{f_\infty}]^\top \in \mathbb{R}^\infty$, $A_s = \text{diag}\{\lambda_1, \cdots, \lambda_m\}$, $f_s = \langle \varphi_s, f \rangle$, $\phi_s^k = \langle \varphi_s, \phi^k \rangle$, $x_{s_0} = \langle \varphi_s, \chi_0 \rangle$, $A_f = \text{diag}\{\lambda_{m+1}, \cdots, \lambda_\infty\}$, $f_f = \langle \varphi_f, f \rangle$, $\phi_f^k = \langle \varphi_f, \phi^k \rangle$, $\chi_{f_0} = \langle \varphi_f, \chi_0 \rangle$ with $\varphi_s = [\varphi_1, \cdots, \varphi_m]^\top$ and $\varphi_f = [\varphi_{m+1}, \cdots, \varphi_\infty]^\top$. By neglecting the fast modes, we can obtain the following finite-dimensional ODE model to characterize the dominant dynamics of the PDE system in (3):

$$\dot{x}_s = A_s x_s + f_s(x_s, u) + \beta(t-t_0)\phi_s^k(x_s, u),\ x_s(0) = x_{s_0}. \tag{5}$$

Remark 1. *The process of model-reduction in Eqs. (3)–(5) is based on the Galerkin method, as adopted in existing works [6], [19], which is included here for the completeness of presentation. Note that such a process is not the major contribution of this paper, thus its thorough analysis is not provided here. Interested readers are referred to [34], [6], [19] for more details.*

In the next sections, a novel FDI scheme will be proposed based on the ODE system (5), so as to achieve accurate detection and isolation for the occurring fault $\phi_s^k(x_s, u)$. Note that since the functions $f_s(x_s, u)$ and $\phi_s^k(x_s, u)$ are both unknown, the model (5) cannot be directly used for the FDI design. In view of this, the FDI scheme proposed in this paper will consist of three components: (i) adaptive dynamics learning, to achieve locally-accurate identification of the uncertain dynamics $f_s(x_s, u)$ and $\phi_s^k(x_s, u)$ in system (5) under the normal mode and all faulty modes; (ii) FD scheme, to achieve rapid detection of fault occurrence; and (iii) FI scheme, to realize accurate fault isolation, which will be activated once the occurring fault is detected.

III. Identification of System Uncertain Dynamics

In this section, a DL-based adaptive dynamics learning approach will be developed to achieve accurate identification of the uncertain dynamics $f_s(x_s, u)$ and $\phi_s^k(x_s, u)$ in system (5) under all normal and faulty modes.

Consider the following faulty dynamic systems:

$$\dot{x}_s = A_s x_s + f_s(x_s, u) + \phi_s^k(x_s, u), \tag{6}$$

where $k = 0, 1, \cdots, N$ denotes the k-th faulty mode, with $k = 0$ representing the normal mode, i.e., $\phi_s^0(x_s, u) \equiv 0$. Since the system uncertainty $f_s(x_s, u)$ and occurring fault $\phi_s^k(x_s, u)$ in (6) cannot be decoupled, by considering them together and defining a general fault function $\eta^k(x_s, u) := f_s(x_s, u) + \phi_s^k(x_s, u)$, we can rewrite the system (6) as:

$$\dot{x}_{s_i} = \lambda_i x_{s_i} + \eta_i^k(x_s, u), \quad i = 1, 2, \cdots, m. \tag{7}$$

For the unknown function $\eta_i^k(x_s, u)$ in (7), according to the RBF NN approximation theory as presented in Section II-A, we know that there exists an ideal constant NN weight vector $W_i^{k*} \in \mathbb{R}^{N_n}$ (with N_n denoting the number of NN nodes) such that

$$\eta_i^k(x_s, u) = W_i^{k*\top} S(x_s, u) + \varepsilon_{i_0}^k, \tag{8}$$

where $S(x_s, u) : \mathbb{R}^m \times \mathbb{R}^q \to \mathbb{R}^{N_n}$ is a smooth RBF vector and $\varepsilon_{i_0}^k$ is the estimation error satisfying $|\varepsilon_{i_0}^k| < \varepsilon_i^*$ with ε_i^* being a positive constant that can be made arbitrarily small given a sufficiently large number of neurons. Based on this, an adaptive dynamics identifier can be constructed:

$$\begin{aligned} \dot{\hat{x}}_i &= -a_i(\hat{x}_i - x_{s_i}) + \lambda_i x_{s_i} + \hat{W}_i^{k\top} S(x_s, u), \\ \dot{\hat{W}}_i^k &= -\sigma_i \Gamma_i \hat{W}_i^k - \Gamma_i (\hat{x}_i - x_{s_i}) S(x_s, u), \end{aligned} \tag{9}$$

for all $i = 1, 2, \cdots, m$ and $k = 0, 1, \cdots, N$, where $\hat{x}_i$ is the identifier state, x_{s_i} is the state of system (7), $\hat{W}_i^k \in \mathbb{R}^{N_n}$ is the estimate of W_i^{k*} in (8), $a_i > 0$, $\Gamma_i = \Gamma_i^\top > 0$, $\sigma_i > 0$ are design constants with σ_i being a small number.

Theorem 1. *Consider the adaptive learning system consisting of the plant (7) and the identifier (9). Under Assumption 1 with initial condition $\hat{W}_i^k(0) = 0$, for all $i = 1, \cdots, m$ and $k = 0, 1, \cdots, N$, we have: (i) all signals in the system remain bounded; (ii) the estimation error $|\hat{x}_i - x_{s_i}|$ converges to a small neighborhood around the origin; and (iii) a locally-accurate approximation of the unknown function $\eta_i^k(x_s, u)$ is achieved by $\hat{W}_i^{k\top} S(x_s, u)$ as well as $\bar{W}_i^{k\top} S(x_s, u)$ along the recurrent system trajectory (x_s, u), where $\bar{W}_i^k := \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \hat{W}_i^k(\tau) d\tau$ with $[t_1, t_2]$ representing a time segment after the transient process.*

Detailed proof can be completed by following a similar line of the proof of [23, Th. 3.1], thus is omitted here.

Remark 2. *Implementing (9) requires information of the system state $x_s(t)$, which can be obtained by measuring the state signal $x(z, t)$ from the original PDE system (1)–(2) via $x_s(t) = \langle \varphi_s(z), x(z, t) \rangle$.*

Through the above learning process, the knowledge of unknown function $\eta_i^k(x_s, u)$ of (7) can finally be obtained and stored in the constant RBF NN model $\bar{W}_i^{k\top} S(x_s, u)$, i.e.,

$$\eta_i^k(x_s, u) = \bar{W}_i^{k\top} S(x_s, u) + \varepsilon_i^k, \tag{10}$$

for all $i = 1, 2, \cdots, m$ and $k = 0, 1, \cdots, N$, where the approximation error ε_i^k satisfies $|\varepsilon_i^k| = O(\varepsilon_i^*) < \xi_i^*$, with ξ_i^* being a positive constant that can be made arbitrarily small by constructing a sufficiently large number of neurons [22].

IV. Fault Detection Scheme

With the results obtained from the above section, a novel FD scheme will be proposed in this section to achieve rapid FD of system (5). The associated analysis of FD performance will also be provided.

A. FD Estimator Design and Decision Making

With the constant RBF NN models $\bar{W}_i^{0\top} S(x_s, u)$ in (10), a bank of FD estimators can be constructed as follows:

$$\dot{\bar{x}}_i^0 = -b_i^0(\bar{x}_i^0 - x_{s_i}) + \lambda_i x_{s_i} + \bar{W}_i^{0\top} S(x_s, u), \tag{11}$$

where $i = 1, \cdots, m$, $\bar{x}_i^0$ is the estimator state with initial condition $\bar{x}_i^0(0) = x_{s_i}(0)$, x_{s_i} is the i-th state of system (5), b_i^0 is a positive design constant, λ_i is the i-th diagonal element of A_s in (5), and $\bar{W}_i^{0\top} S(x_s, u)$ is used to approximate the function $f_{s_i}(x_s, u)$ in (5). Comparing the FD estimators (11) with the monitored system (5), and based on (10), the following residual system (with residual $\tilde{x}_i^0 := \bar{x}_i^0 - x_{s_i}$) can be derived:

$$\dot{\tilde{x}}_i^0 = -b_i^0 \tilde{x}_i^0 - \varepsilon_i^0 - \beta(t - t_0)\phi_{s_i}^k(x_s, u), \tag{12}$$

where ε_i^0 is the approximation error of model $\bar{W}_i^{0\top} S(x_s, u)$ for function $f_{s_i}(x_s, u)$ as defined in (10), and $\phi_{s_i}^k(x_s, u)$ is the faulty dynamics occurring in system (5). The $\mathcal{L}_1$ norm of residual signal $\tilde{x}_i^0$ in (12), i.e., $\left\|\tilde{x}_i^0(t)\right\|_1 = \frac{1}{T}\int_{t-T}^{t} \left|\tilde{x}_i^0(\tau)\right| d\tau$ $(t > T)$ with T being a design parameter, will be used for real-time FD decision making. Before proceeding further, a threshold, denoted as $\bar{e}_i^0$, will be further designed to upper bound $\left\|\tilde{x}_i^0(t)\right\|_1$ when the monitored system (5) is operating in normal mode (i.e., for time $t < t_0$). To this end, consider the residual system (12) for time $t < t_0$, note that $\tilde{x}_i^0(0) = 0$ and $|\varepsilon_i^0| < \xi_i^*$, for all $i = 1, \cdots, m$, the system state $\tilde{x}_i^0$ satisfies:

$$\begin{aligned} \left|\tilde{x}_i^0(t)\right| &= \left|\tilde{x}_i^0(0)e^{-b_i^0 t} - \int_0^t e^{-b_i^0(t-\tau)}\varepsilon_i^0 d\tau\right| \\ &\leq \int_0^t e^{-b_i^0(t-\tau)}\left|\varepsilon_i^0\right| d\tau < \int_0^t e^{-b_i^0(t-\tau)}\xi_i^* d\tau < \frac{\xi_i^*}{b_i^0}. \end{aligned} \tag{13}$$

It implies that the FD residual signal $\left\|\tilde{x}_i^0(t)\right\|_1 < \frac{\xi_i^*}{b_i^0}$ holds under the normal mode for all time $t < t_0$. Based on this, the FD threshold $\bar{e}_i^0$ can be designed as:

$$\bar{e}_i^0 := \frac{1}{b_i^0}(\xi_i^* + \varrho_i), \quad i = 1, \cdots, m, \tag{14}$$

where ξ_i^* is a small constant given in (10), b_i^0 is a design constant from (11), and $\varrho_i \geq 0$ is a small constant added as an auxiliary parameter for preventing possible FD misjudgment.

With the FD estimators (11) and FD thresholds (14), the FD decision making is based on the following principle: when no fault occurs in the monitored system (5), the residuals $\left\|\tilde{x}_i^0(t)\right\|_1$ remain smaller than the corresponding thresholds $\bar{e}_i^0$ for all $i = 1, \cdots, m$. If there exists a time instant t_d, such that, for some $i \in \{1, \cdots, m\}$, the FD residuals $\left\|\tilde{x}_i^0(t_d)\right\|_1$ become larger than the corresponding thresholds $\bar{e}_i^0$, i.e., $\left\|\tilde{x}_i^0(t_d)\right\|_1 > \bar{e}_i^0$, it indicates that a certain fault must occur in the system (5). As a result, the occurrence of fault can be detected at time t_d. The idea is formalized as follows:

Fault detection decision making: Compare the FD residual signals $\left\|\tilde{x}_i^0(t)\right\|_1$ with the FD thresholds $\bar{e}_i^0$ for all $i = 1, \cdots, m$. If there exists a finite time t_d, such that, for some $i \in \{1, \cdots, m\}$, $\left\|\tilde{x}_i^0(t_d)\right\|_1 > \bar{e}_i^0$ holds. Then, the occurrence of a fault is deduced at time t_d.

Remark 3. *The parameter ξ_i^* in (14) represents the upper bound of steady absolute approximation error $\left|\bar{W}_i^{k\top} S(x_s, u) - \eta_i^k(x_s, u)\right|$ of (10) for all $k = 0, 1, \cdots, m$. Direct derivation of this parameter is quite difficult since the function $\eta_i^k(x_s, u)$ is not available. Alternatively, the value of ξ_i^* could be evaluated in the following way: in the training phase of Section III, with the obtained constant models $\bar{W}_i^{k\top} S(x_s, u)$, a bank of estimators in the form of (11) (with $k = 0, 1, \cdots, N$) can be developed by setting $b_i^k = 1$. Then, following a similar line of the analysis in Eqs. (11)–(13), it can be proved that the associated state error $\tilde{x}_i^k = \bar{x}_i^k - x_{s_i}$ satisfies: $|\tilde{x}_i^k(t)| < \xi_i^*$. Thus, with such estimators, the value of ξ_i^* can be obtained as the upper bound of steady absolute state error $|\tilde{x}_i^k|$, $\forall k = 0, 1, \cdots, N$.*

Remark 4. *The parameter ϱ_i in the FD threshold (14) is designed to improve robustness against system uncertainties. More specific, note that although our FD scheme is developed based on the approximate ODE system (5), the real-time FD process will be carried out on the original PDE system (4). The associated system dynamics of models (5) and (4) have a small difference due to the fast dynamics of state χ_f. The parameter ϱ_i is thus introduced to compensate such a difference and mitigate its potential effects on the FD performance.*

Remark 5. *The FD threshold (14) can be made very small, because the parameter ξ_i^* can be made arbitrarily small by constructing a sufficiently large number of neurons in the training process of Section III, and the parameter ϱ_i can be selected also as a very small number.*

Remark 6. *Most of existing FD schemes (e.g., [10], [13], [32], [9]) cannot deal with the effect of system uncertainty on FD process. As a result, these schemes require the occurring faults to be of sufficiently large magnitudes (larger than that of system uncertainty) for successful detection, quite limiting their fault detectability. However, this issue can be addressed under our scheme. Specifically, as established in Section III, the system uncertainty $f_{s_i}(x_s, u)$ in (5) can be accurately identified with the DL-based dynamics learning scheme and the associated knowledge can be obtained and stored in a constant NN model $\bar{W}_i^{0\top} S(x_s, u)$ of (10). By using this model to design the FD estimator (11), the fault dynamics $\phi_{s_i}^k(x_s, u)$ in system (5) can be accurately distinguished from the uncertain dynamics $f_{s_i}(x_s, u)$, and will be captured by the FD signal $\tilde{x}_i^0$ in (12) for accurate detection. It will facilitate our scheme to develop improved fault detectability compared to the ones in [10], [13], [32], [9]. Associated rigorous analysis will be conducted in the next section.*

Remark 7. *Rapid FD process can be achieved with our approach. Note that the FD estimators of (11) are designed with the constant NN models $\bar{W}_i^{0\top} S(x_s, u)$, whose implementation does not involve any parameter adaptation. This will largely shorten the FD time, such that FD process of (11) can be achieved in a rapid manner.*

B. Detectability Condition

To analyze the performance of the proposed FD scheme, in the following, we will study the fault detectability condition,

i.e., under what conditions the occurring fault in system (5) is detectable with our proposed FD scheme.

Theorem 2. *Consider the system (5) and the fault detection system consisting of estimators (11) and thresholds (14). If there exists a time interval $I = [t_a, t_b] \subseteq [t_b - T, t_b]$ with $t_a \geq t_0$, such that for some $i \in \{1, \cdots, m\}$,*

$$\left|\phi_{s_i}^k(x_s(t), u(t))\right| > 2\xi_i^* + 2\varrho_i, \quad \forall t \in I \tag{15}$$

and

$$l := t_b - t_a \geq \frac{1}{b_i^0}\ln\frac{7\mu_i - 6\xi_i^*}{\mu_i - 2\xi_i^*} + \frac{T(4\xi_i^* + 4\varrho_i)}{3\mu_i - 2\xi_i^*}, \tag{16}$$

where $\mu_i := \min\{\left|\phi_{s_i}^k(x_s, u)\right|, \forall t \in I\}$, then, $\left\|\tilde{x}_i^0(t_b)\right\|_1 > \bar{e}_i^0$ holds and the occurrence of a fault will be detected at time t_b, i.e., $t_d = t_b$.

Proof. Consider the residual signal $\tilde{x}_i^0(t)$ of (12). In the time interval I, we assume that there exists a subinterval $I' \subseteq I$ such that the signal $\left|\tilde{x}_i^0(t)\right|$ has a very small magnitude, i.e.,

$$I' := \left\{t \in I : \left|\tilde{x}_i^0(t)\right| \leq \frac{3\mu_i - 2\xi_i^*}{4b_i^0}\right\}, \tag{17}$$

where $\mu_i > 2\xi_i^* + 2\varrho_i$ from (15). For $t \in I'$, by denoting $t_a' = \min\{t, t \in I'\}$, the residual signal $\tilde{x}_i^0$ of (12) satisfies

$$\begin{aligned}\left|\tilde{x}_i^0\right| &= \left|\tilde{x}_i^0(t_a')e^{-b_i^0(t-t_a')} - \int_{t_a'}^t e^{-b_i^0(t-\tau)}(\phi_{s_i}^k(x_s, u) + \varepsilon_i^0)d\tau\right| \\ &\geq \left|\int_{t_a'}^t e^{-b_i^0(t-\tau)}(\phi_{s_i}^k(x_s, u) + \varepsilon_i^0)d\tau\right| - \left|\tilde{x}_i^0(t_a')\right| e^{-b_i^0(t-t_a')}.\end{aligned} \tag{18}$$

From (15) and (10), for all $t \in I'$, $\phi_{s_i}^k(x_s, u) + \varepsilon_i^0$ satisfies

$$\left|\phi_{s_i}^k(x_s, u) + \varepsilon_i^0\right| \geq \left|\phi_{s_i}^k(x_s, u)\right| - \left|\varepsilon_i^0\right| \geq \mu_i - \xi_i^*. \tag{19}$$

Note that $\mu_i - \xi_i^* > 0$, it is easily seen that $\phi_{s_i}^k(x_s, u) + \varepsilon_i^0$ has an unchanged sign for all $t \in I'$, such that

$$\begin{aligned}&\left|\int_{t_a'}^t e^{-b_i^0(t-\tau)}(\phi_{s_i}^k(x_s, u) + \varepsilon_i^0)d\tau\right| \\ &= \int_{t_a'}^t e^{-b_i^0(t-\tau)}\left|\phi_{s_i}^k(x_s, u) + \varepsilon_i^0\right|d\tau \\ &\geq \int_{t_a'}^t e^{-b_i^0(t-\tau)}(\mu_i - \xi_i^*)d\tau = \frac{\mu_i - \xi_i^*}{b_i^0}(1 - e^{-b_i^0(t-t_a')}).\end{aligned} \tag{20}$$

Then, since $\left|\tilde{x}_i^0(t_a')\right| \leq \frac{3\mu_i - 2\xi_i^*}{4b_i^0}$, inequality (18) reduces to

$$\left|\tilde{x}_i^0(t)\right| \geq \frac{\mu_i - \xi_i^*}{b_i^0}(1 - e^{-b_i^0(t-t_a')}) - \frac{3\mu_i - 2\xi_i^*}{4b_i^0}e^{-b_i^0(t-t_a')}. \tag{21}$$

As a result, it can be deduced that $\left|\tilde{x}_i^0(t)\right| > \frac{3\mu_i - 2\xi_i^*}{4b_i^0}$ holds for $t - t_a' > \frac{1}{b_i^0}\ln\frac{7\mu_i - 6\xi_i^*}{\mu_i - 2\xi_i^*}$, and the length of time interval I' in (17), denoted by l', satisfies $l' < \frac{1}{b_i^0}\ln\frac{7\mu_i - 6\xi_i^*}{\mu_i - 2\xi_i^*}$. Furthermore, it is easily verified that there exists at most one subinterval I' defined in (17) over the time interval I. This implies that for the time interval $I - I'$, we have:

$$\left|\tilde{x}_i^0(t)\right| > \frac{3\mu_i - 2\xi_i^*}{4b_i^0}, \quad \forall t \in I - I', \tag{22}$$

and the length of time interval $I - I'$ (i.e., $l - l'$) satisfies $l - l' > l - \frac{1}{b_i^0}\ln\frac{7\mu_i - 6\xi_i^*}{\mu_i - 2\xi_i^*}$. Based on this, from (16), we have:

$$\begin{aligned}\left\|\tilde{x}_i^0(t_b)\right\|_1 &= \frac{1}{T}\int_{t_b - T}^{t_b}\left|\tilde{x}_i^0(\tau)\right|d\tau \geq \frac{1}{T}\int_{I - I'}\left|\tilde{x}_i^0(\tau)\right|d\tau \\ &> \frac{1}{T}\int_{I - I'}\frac{3\mu_i - 2\xi_i^*}{4b_i^0}d\tau = \frac{1}{T}(l - l')\frac{3\mu_i - 2\xi_i^*}{4b_i^0} \\ &> \frac{1}{T}(l - \frac{1}{b_i^0}\ln\frac{7\mu_i - 6\xi_i^*}{\mu_i - 2\xi_i^*})\frac{3\mu_i - 2\xi_i^*}{4b_i^0} \geq \frac{\xi_i^* + \varrho_i}{b_i^0}.\end{aligned} \tag{23}$$

Thus, $\left\|\tilde{x}_i^0(t_b)\right\|_1 > \bar{e}_i^0$ holds and the occurrence of fault in system (5) can be detected at time t_b. This ends the proof. □

Remark 8. *The detectability conditions (15)–(16) show that if there exists a time interval $[t_a, t_b]$ of (16) such that the occurring fault $\phi_{s_i}^k(x_s, u)$ has a sufficiently large magnitude, i.e., larger than the lower bound $2\xi_i^* + 2\varrho_i$ of (15), then, fault detection can be achieved. Particularly, note that the lower bound $2\xi_i^* + 2\varrho_i$ can be made arbitrarily small, as argued in Remark 5, the conditions (15)–(16) are thus satisfiable even for those faults with relatively small magnitudes.*

V. Fault Isolation Scheme

Once the occurring fault is detected at time t_d ($t_d > t_0$), the FI scheme will be activated to identify the type of the occurring fault. This section will present the design of such an FI scheme, as well as the associated analysis of FI performance. To ease the presentation, we assume without loss of generality that an unknown fault l' that is similar to (but not necessarily perfectly match) the trained fault l ($l \in \{1, \cdots, N\}$) is occurring in system (5), i.e.,

$$\dot{x}_s = A_s x_s + f_s(x_s, u) + \phi_s^{l'}(x_s, u). \tag{24}$$

A. FI Estimator Design and Decision Making

With the constant models $\bar{W}_i^{k\top}S(x_s, u)$ of (10) obtained from the identification phase of Section III, we propose to construct a bank of FI estimators in the following form

$$\dot{\bar{x}}_i^k = -b_i(\bar{x}_i^k - x_{s_i}) + \lambda_i x_{s_i} + \bar{W}_i^{k\top}S(x_s, u), \tag{25}$$

where $i = 1, \cdots, m$, $k = 1, \cdots, N$, $\bar{x}_i^k$ is the estimator state with initial condition $\bar{x}_i^k(t_d) = x_{s_i}(t_d)$, x_{s_i} is the i-th state of system (24), λ_i is the i-th diagonal element of matrix A_s in (24), b_i is a positive design constant, and $\bar{W}_i^{k\top}S(x_s, u)$ approximates the function $\eta_i^k(x_s, u) = f_{s_i}(x_s, u) + \phi_{s_i}^k(x_s, u)$ of system (6). Comparing the FI estimators (25) with the monitored system (24), and based on (10), the residual systems (with residual $\tilde{x}_i^k := \bar{x}_i^k - x_{s_i}$) can be derived as follows:

$$\dot{\tilde{x}}_i^k = -b_i\tilde{x}_i^k - \varepsilon_i^k + \phi_{s_i}^k(x_s, u) - \phi_{s_i}^{l'}(x_s, u), \tag{26}$$

where ε_i^k is the approximation error of model $\bar{W}_i^{k\top}S(x_s, u)$ for function $\eta_i^k(x_s, u)$ as defined in (10), $\phi_{s_i}^k(x_s, u)$ is the k-th faulty dynamics that has been learned/trained in Section III, and $\phi_{s_i}^{l'}(x_s, u)$ is the faulty dynamics occurring in system (24). For the purpose of analysis, we introduce a so-called fault mismatch function $\rho_i^{k,l'}(x_s, u) := \phi_{s_i}^k(x_s, u) - \phi_{s_i}^{l'}(x_s, u)$ to represent the dynamics difference between the trained fault k

and occurring fault l'. Then, the residual system (26) can be rewritten as:

$$\dot{\tilde{x}}_i^k = -b_i\tilde{x}_i^k - \varepsilon_i^k + \rho_i^{k,l'}(x_s,u). \tag{27}$$

Similar to the FD case, the $\mathcal{L}_1$ norm of residual signal $\tilde{x}_i^k$ in (27), i.e., $\left\|\tilde{x}_i^k(t)\right\|_1 = \frac{1}{T}\int_{t-T}^{t}|\tilde{x}_i^k(\tau)|d\tau$ with T being a design parameter, will be utilized for real-time FI decision making.

In the following, for FI decision making, an adaptive threshold, denoted as $\bar{e}_i^k(t)$, will be further designed to upper bound the residual signal $\left\|\tilde{x}_i^k(t)\right\|_1$ when the occurring fault l' in (24) is similar to the trained fault l. To this end, the following assumption on the dynamics difference between the occurring fault l' and similar fault l is made.

Assumption 3. *The dynamics difference between any pair of the occurring fault l' and its similar fault l ($l \in \{1,\cdots,N\}$), denoted by $\rho_i^{l,l'}(x_s,u)$, is bounded by a known function $\bar{\rho}_i^l(x_s,u)$, i.e., $\left|\rho_i^{l,l'}(x_s,u)\right| \le \bar{\rho}_i^l(x_s,u)$ for all $i=1,\cdots,m$.*

Remark 9. *Assumption 3 indicates that the occurring fault l' is allowed to have a certain degree of difference from its similar fault l, and such difference can be quantified by the function $\bar{\rho}_i^l(x_s,u)$. In other words, it allows that the occurring fault is not necessarily required to exactly match any of the pre-defined/pre-trained faults, which however is typically required by existing methods of [25], [26], [27]. This property renders our FI scheme a better robust capability of preventing false/missed FI alarm in the presence of slight fault difference during the FI process.*

Based on the above setup, to design the FI adaptive threshold, we consider the l-th residual system in (27), its time-domain solution can be derived as:

$$\tilde{x}_i^l(t) = \tilde{x}_i^l(t_d)e^{-b_i(t-t_d)} + \int_{t_d}^{t} e^{-b_i(t-\tau)}(\rho_i^{l,l'}(x_s,u) - \varepsilon_i^l)d\tau. \tag{28}$$

Note that $\tilde{x}_i^l(t_d) = 0$ and $\left|\varepsilon_i^l\right| < \xi_i^*$ from (10), under Assumption 3, we have:

$$\begin{aligned}\left|\tilde{x}_i^l(t)\right| &\le \int_{t_d}^{t} e^{-b_i(t-\tau)}(\left|\varepsilon_i^l\right| + \left|\rho_i^{l,l'}(x_s,u)\right|)d\tau \\ &< \frac{\xi_i^*}{b_i} + \int_{t_d}^{t} e^{-b_i(t-\tau)}\bar{\rho}_i^l(x_s,u)d\tau.\end{aligned} \tag{29}$$

It guarantees that the FI residual signal $\left\|\tilde{x}_i^l(t)\right\|_1$ satisfies:

$$\left\|\tilde{x}_i^l(t)\right\|_1 < \frac{\xi_i^*}{b_i} + \left\|\int_{t_d}^{t} e^{-b_i(t-\tau)}\bar{\rho}_i^l(x_s,u)d\tau\right\|_1. \tag{30}$$

Thus, the FI adaptive threshold $\bar{e}_i^l(t)$ can be designed as:

$$\bar{e}_i^l(t) := \frac{\xi_i^*}{b_i} + \left\|\int_{t_d}^{t} e^{-b_i(t-\tau)}\bar{\rho}_i^l(x_s,u)d\tau\right\|_1, \tag{31}$$

for all $i=1,\cdots,m$, where ξ_i^* is a small constant given in (10), b_i is a design constant from (25), and $\bar{\rho}_i^l(x_s,u)$ is a known function defined in Assumption 3.

Remark 10. *The FI thresholds (31) can be implemented in a simplified form if the function $\bar{\rho}_i^l(x_s,u)$ is a constant $\bar{\rho}_i^l$. Specifically, with $\bar{\rho}_i^l(x_s,u) = \bar{\rho}_i^l$, the threshold (31) is given by $\bar{e}_i^l(t) = \frac{\xi_i^*}{b_i} + \left\|\frac{\bar{\rho}_i^l}{b_i}(1-e^{-b_i(t-t_d)})\right\|_1$, which can be further simplified as a constant threshold $\bar{e}_i^l = \frac{1}{b_i}(\xi_i^* + \bar{\rho}_i^l)$.*

Consequently, for the monitored system (24), the proposed FI scheme consists of the FI estimators (25) and the adaptive threshold (31). Real-time FI decision making is based on the following principle. If there exists a unique residual system in (27), say the l-th one, such that for all $i=1,\cdots,m$ the residual signals $\tilde{x}_i^l$ satisfy $\left\|\tilde{x}_i^l(t)\right\|_1 \le \bar{e}_i^l(t)$ for all time $t > t_d$, then, it can be deduced that the occurring fault l' in (24) is similar to the trained fault l. Using this idea, the FI decision making scheme can be devised as follows.

Fault isolation decision making: Compare the FI residual signals $\left\|\tilde{x}_i^k(t)\right\|_1$ with the FI adaptive thresholds $\bar{e}_i^k(t)$ for time $t \ge t_d$ and all $i=1,\cdots,m$, $k=1,\cdots,N$. If there exists a unique $l \in \{1,\cdots,N\}$ such that: (i) $\forall i=1,\cdots,m$, $\left\|\tilde{x}_i^l(t)\right\|_1 \le \bar{e}_i^l(t)$ holds for all time $t \ge t_d$; and (ii) $\forall k \in \{1,\cdots,N\}/\{l\}$, $\exists i \in \{1,\cdots,m\}$, $\left\|\tilde{x}_i^k(t^k)\right\|_1 > \bar{e}_i^k(t^k)$ holds at some time instant $t^k > t_d$. Then, the occurring fault l' can be identified similar to the fault l, and the isolation time can be obtained as: $t_{iso} = \max\left\{t^k, k \in \{1,\cdots,N\}/\{l\}\right\}$.

Remark 11. *Similar to the FD scheme of Section IV, the proposed FI scheme can effectively deal with the effect of system uncertainty $f_s(x_s,u)$ in (24) for accurate isolation, and the associated FI process can be achieved in a rapid manner. This is owing to the utilization of constant NN models $\bar{W}_i^{k\top}S(x_s,u)$ (obtained through the training process of Section III) in the design of FI estimators (25).*

Remark 12. *Existing FI schemes in [15], [16], [17] rely on "constant" thresholds for FI, which would limit the ability to separate the temporal dynamics of different type of faults for accurate isolation. These schemes can be applicable only to the faults that have sufficiently distinct differences. For example, the FI schemes in [15], [17] can distinguish the actuator faults occurring at different locations, but cannot recognize the actuator faults that occur at the same location but have different magnitudes. Advanced over these schemes, our approach design an "adaptive" threshold of (31) by using the nonlinear function $\bar{\rho}_i^l(x_s,u)$ that can accurately specify the similarity of each l-th type of faults, as defined in Assumption 3. With such a threshold, our FI scheme can achieve accurate isolation even for the faults that have relatively small differences. Improved FI accuracy and fault isolatability with our scheme compared to the ones in [15], [16], [17] will be demonstrated later.*

Remark 13. *Our FI scheme provides a unified framework to address the isolation problem of general faults occurring in the PDE system (1)–(2), which possesses enhanced applicability compared to most existing FI schemes, e.g., [6], [15], [16], [17], [18]. Specifically, note that the schemes of [6], [15], [16], [17], [18] are applicable only to some special cases. For example, the FI scheme in [6] is tailored only to systems with precisely known model; the method of [16] is tailored to linear PDE systems; while the approaches in [15], [17], [18] are applicable only for actuator faults. Advanced over these approaches, our FI scheme is developed for a class*

of general PDE system with the form of (1)–(2), in which the occurring fault $\phi^k(x,u)$ is not required to be of any special type and the system model is allowed to have uncertain nonlinear component $f(x,u)$.

B. Isolatability Condition

To analyze the performance of the proposed FI scheme, the fault isolatability condition will be studied, i.e., under what conditions the occurring fault l' in system (24) can be identified similar to a unique trained fault l.

Theorem 3. *Consider the monitored system (24) and the fault isolation system consisting of estimators (25) and adaptive thresholds (31). For each $k \in \{1,\cdots,N\}/\{l\}$ and some $i \in \{1,\cdots,m\}$, if there exists a time interval $I^k = [t_a^k, t_b^k] \subseteq [t_b^k - T, t_b^k]$ with $t_a^k \geq t_d$, such that*

$$\left|\rho_i^{k,l'}(x_s,u)\right| > \bar{\rho}_i^k(x_s,u) + 2\xi_i^*, \quad \forall t \in I^k, \tag{32}$$

and

$$l^k := t_a^k - t_b^k \geq \frac{2\xi_i^* + 2\bar{\rho}_{i_{\max}}^k}{\mu_i + 2\bar{\rho}_{i_{\max}}^k}\left(T + \frac{1}{b_i}\ln\frac{\mu_i + 2\bar{\rho}_{i_{\max}}^k + \xi_i^*}{\mu_i - 2\xi_i^*}\right) + \frac{\mu_i - 2\xi_i^*}{\mu_i + 2\bar{\rho}_{i_{\max}}^k}\left(\frac{1}{b_i}\ln\frac{3\mu_i + 4\bar{\rho}_{i_{\max}}^k}{\mu_i - 2\xi_i^*}\right), \tag{33}$$

where $\mu_i := \min\{\left|\rho_i^{k,l'}(x_s,u)\right| - \bar{\rho}_i^k(x_s,u), \forall t \in I^k\}$, and $\bar{\rho}_{i_{\max}}^k := \max\{\bar{\rho}_i^k(x_s,u), \forall t \geq t_d\}$, then, $\left\|\tilde{x}_i^k(t_b^k)\right\|_1 > \bar{e}_i^k(t_b^k)$ holds, the occurring fault l' will be identified similar to fault l, and the isolation time is obtained as $t_{iso} = \max\left\{t_b^k, \forall k \in \{1,\cdots,N\}/\{l\}\right\}$.

Proof. Consider the k-th residual signal $\tilde{x}_i^k(t)$ of (27) and the associated FI threshold $\bar{e}_i^k(t)$ of (31), with $k \in \{1,\cdots,N\}/\{l\}$. For the purpose of analysis, we introduce a new variable $\vartheta_i^k(t)$ satisfying

$$\vartheta_i^k(t) := \frac{\xi_i^*}{b_i} + \int_{t_d}^t e^{-b_i(t-\tau)}\bar{\rho}_i^k(x_s,u)d\tau, \quad t \geq t_d, \tag{34}$$

and thus we have

$$\vartheta_i^k(t) \leq \frac{\xi_i^*}{b_i} + \int_{t_d}^t e^{-b_i(t-\tau)}\bar{\rho}_{i_{\max}}^k d\tau < \frac{\xi_i^* + \bar{\rho}_{i_{\max}}^k}{b_i}, \tag{35}$$

where $\bar{\rho}_{i_{\max}}^k = \max\{\bar{\rho}_i^k(x_s,u), \forall t \geq t_d\}$. Obviously, to guarantee that $\left\|\tilde{x}_i^k(t)\right\|_1 > \bar{e}_i^k(t)$ holds at a time $t = t_b^k$, in light of definitions (31) and (34), it is necessary to examine the magnitude of $\left|\tilde{x}_i^k(t)\right| - \vartheta_i^k(t)$ for $t \in [t_b^k - T, t_b^k]$.

We first consider the time interval $I^k \subseteq [t_b^k - T, t_b^k]$. Assume that there exists a subinterval $I_1^k \subseteq I^k$ such that the residual signal $\left|\tilde{x}_i^k(t)\right| - \vartheta_i^k(t)$ has a very small magnitude, i.e.,

$$I_1^k := \left\{t \in I^k : \left|\tilde{x}_i^k(t)\right| - \vartheta_i^k(t) \leq \frac{\mu_i - 2\xi_i^*}{2b_i}\right\}, \tag{36}$$

where $\mu_i = \min\{\left|\rho_i^{k,l'}(x_s,u)\right| - \bar{\rho}_i^k(x_s,u), \forall t \in I^k\} > 2\xi_i^*$ from (32). For $t \in I_1^k$, by denoting $t_a^{k'} = \min\{t, t \in I_1^k\}$, based on (10), the residual signal $\tilde{x}_i^k(t)$ of (27) satisfies

$$\begin{aligned}
&\left|\tilde{x}_i^k(t)\right| \\
&= \left|\tilde{x}_i^k(t_a^{k'})e^{-b_i(t-t_a^{k'})} + \int_{t_a^{k'}}^t e^{-b_i(t-\tau)}(\rho_i^{k,l'}(x_s,u) - \varepsilon_i^k)d\tau\right| \\
&\geq \left|\int_{t_a^{k'}}^t e^{-b_i(t-\tau)}\rho_i^{k,l'}(x_s,u)d\tau\right| - \int_{t_a^{k'}}^t e^{-b_i(t-\tau)}\left|\varepsilon_i^k\right|d\tau \\
&\quad - \left|\tilde{x}_i^k(t_a^{k'})\right|e^{-b_i(t-t_a^{k'})} \\
&> \left|\int_{t_a^{k'}}^t e^{-b_i(t-\tau)}\rho_i^{k,l'}(x_s,u)d\tau\right| - \frac{\xi_i^*}{b_i} - \left|\tilde{x}_i^k(t_a^{k'})\right|e^{-b_i(t-t_a^{k'})}.
\end{aligned} \tag{37}$$

Under condition (32), it is seen that for all $t \in I_1^k$, $\rho_i^{k,l'}(x_s,u)$ has an unchanged sign. Then, inequality (37) yields:

$$\begin{aligned}
\left|\tilde{x}_i^k(t)\right| &> \int_{t_a^{k'}}^t e^{-b_i(t-\tau)}\left|\rho_i^{k,l'}(x_s,u)\right|d\tau - \frac{\xi_i^*}{b_i} \\
&\quad - \left|\tilde{x}_i^k(t_a^{k'})\right|e^{-b_i(t-t_a^{k'})}.
\end{aligned} \tag{38}$$

From (34) and (38), we have:

$$\begin{aligned}
&\left|\tilde{x}_i^k(t)\right| - \vartheta_i^k(t) > \int_{t_a^{k'}}^t e^{-b_i(t-\tau)}\left|\rho_i^{k,l'}(x_s,u)\right|d\tau - \frac{2\xi_i^*}{b_i} \\
&\quad - \int_{t_d}^t e^{-b_i(t-\tau)}\bar{\rho}_i^k(x_s,u)d\tau - \left|\tilde{x}_i^k(t_a^{k'})\right|e^{-b_i(t-t_a^{k'})} \\
&= \int_{t_a^{k'}}^t e^{-b_i(t-\tau)}(\left|\rho_i^{k,l'}(x_s,u)\right| - \bar{\rho}_i^k(x_s,u))d\tau - \frac{2\xi_i^*}{b_i} \\
&\quad - (\vartheta_i^k(t_a^{k'}) - \frac{\xi_i^*}{b_i})e^{-b_i(t-t_a^{k'})} - \left|\tilde{x}_i^k(t_a^{k'})\right|e^{-b_i(t-t_a^{k'})} \\
&\geq \int_{t_a^{k'}}^t e^{-b_i(t-\tau)}\mu_i d\tau - \frac{2\xi_i^*}{b_i} \\
&\quad - (2\vartheta_i^k(t_a^{k'}) + \frac{\mu_i - 4\xi_i^*}{2b_i})e^{-b_i(t-t_a^{k'})} \\
&> \frac{\mu_i}{b_i}(1 - e^{-b_i(t-t_a^{k'})}) - \frac{2\xi_i^*}{b_i} - \frac{4\bar{\rho}_{i_{\max}}^k + \mu_i}{2b_i}e^{-b_i(t-t_a^{k'})},
\end{aligned} \tag{39}$$

where $\left|\tilde{x}_i^k(t_a^{k'})\right| \leq \vartheta_i^k(t_a^{k'}) + \frac{\mu_i - 2\xi_i^*}{2b_i}$, $\left|\rho_i^{k,l'}(x_s,u)\right| - \bar{\rho}_i^k(x_s,u) \geq \mu_i$, and $\vartheta_i^k(t_a^{k'}) < \frac{\xi_i^* + \bar{\rho}_{i_{\max}}^k}{b_i}$, according to (35) and (36). Based on (39), it can be deduced that: $\left|\tilde{x}_i^k(t)\right| - \vartheta_i^k(t) > \frac{\mu_i - 2\xi_i^*}{2b_i}$ holds for $t - t_a^{k'} > \frac{1}{b_i}\ln\frac{3\mu_i + 4\bar{\rho}_{i_{\max}}^k}{\mu_i - 2\xi_i^*}$. Thus, the length of time interval I_1^k in (36), denoted by l_1, satisfies:

$$l_1 \leq \frac{1}{b_i}\ln\frac{3\mu_i + 4\bar{\rho}_{i_{\max}}^k}{\mu_i - 2\xi_i^*}. \tag{40}$$

It is easy to verify that there exists at most one subinterval I_1^k of (36) in the time interval I^k, which implies

$$\left|\tilde{x}_i^k(t)\right| - \vartheta_i^k(t) > \frac{\mu_i - 2\xi_i^*}{2b_i}, \quad \forall t \in I^k - I_1^k. \tag{41}$$

Next, following a similar line of the above analysis, we can further deduce that there exists at most one subinterval I_2^k in the time interval I_1^k, such that:

$$\begin{aligned} &\left|\tilde{x}_i^k(t)\right| - \vartheta_i^k(t) \le 0, \quad \forall t \in I_2^k; \\ 0 < &\left|\tilde{x}_i^k(t)\right| - \vartheta_i^k(t) \le \frac{\mu_i - 2\xi_i^*}{2b_i}, \quad \forall t \in I_1^k - I_2^k, \end{aligned} \tag{42}$$

and the length of the time interval I_2^k, denoted by l_2, satisfies:

$$l_2 \le \frac{1}{b_i} \ln \frac{\mu_i + 2\bar{\rho}_{i_{\max}}^k + \xi_i^*}{\mu_i - 2\xi_i^*}. \tag{43}$$

Particularly, note that $\vartheta_i^k(t) < \frac{\xi_i^* + \bar{\rho}_{i_{\max}}^k}{b_i}$ from (35), inequality (42) yields:

$$-\frac{\xi_i^* + \bar{\rho}_{i_{\max}}^k}{b_i} < \left|\tilde{x}_i^k(t)\right| - \vartheta_i^k(t) \le 0, \quad \forall t \in I_2^k. \tag{44}$$

Finally, consider the signal $\left|\tilde{x}_i^k(t)\right| - \vartheta_i^k(t)$ in the time interval $[t_b^k - T, t_b^k]$. Dividing $[t_b^k - T, t_b^k]$ into four subintervals, i.e., $[t_b^k - T, t_b^k] = ([t_b^k - T, t_b^k] - I^k) \cup (I^k - I_1^k) \cup (I_1^k - I_2^k) \cup I_2^k$, from (35) and (40)–(44), we obtain:

$$\begin{aligned} &\int_{[t_b^k - T, t_b^k] - I^k} (\left|\tilde{x}_i^k(\tau)\right| - \vartheta_i^k(\tau)) d\tau \\ &> \int_{[t_b^k - T, t_b^k] - I^k} (-\frac{\xi_i^* + \bar{\rho}_{i_{\max}}^k}{b_i}) d\tau = -(T - l^k) \frac{\xi_i^* + \bar{\rho}_{i_{\max}}^k}{b_i}; \\ &\int_{I^k - I_1^k} (\left|\tilde{x}_i^k(\tau)\right| - \vartheta_i^k(\tau)) d\tau > \int_{I^k - I_1^k} \frac{\mu_i - 2\xi_i^*}{2b_i} d\tau \\ &= (l^k - l_1) \frac{\mu_i - 2\xi_i^*}{2b_i} > (l^k - \frac{1}{b_i} \ln \frac{3\mu_i + 4\bar{\rho}_{i_{\max}}^k}{\mu_i - 2\xi_i^*}) \frac{\mu_i - 2\xi_i^*}{2b_i}; \\ &\int_{I_1^k - I_2^k} (\left|\tilde{x}_i^k(\tau)\right| - \vartheta_i^k(\tau)) d\tau > (l_1 - l_2) 0 = 0; \\ &\int_{I_2^k} (\left|\tilde{x}_i^k(\tau)\right| - \vartheta_i^k(\tau)) d\tau > \int_{I_2^k} (-\frac{\xi_i^* + \bar{\rho}_{i_{\max}}^k}{b_i}) d\tau \\ &= -l_2 \frac{\xi_i^* + \bar{\rho}_{i_{\max}}^k}{b_i} > -(\frac{1}{b_i} \ln \frac{\mu_i + 2\bar{\rho}_{i_{\max}}^k + \xi_i^*}{\mu_i - 2\xi_i^*}) \frac{\xi_i^* + \bar{\rho}_{i_{\max}}^k}{b_i}. \end{aligned} \tag{45}$$

Based on this, with condition (33), we have:

$$\begin{aligned} &\left\|\tilde{x}_i^k(t_b^k)\right\|_1 - \bar{e}_i^k(t_b^k) \\ =&\frac{1}{T} \int_{([t_b^k - T, t_b^k] - I^k) \cup (I^k - I_1^k) \cup (I_1^k - I_2^k) \cup I_2^k} (\left|\tilde{x}_i^k(\tau)\right| - \vartheta_i^k(\tau)) d\tau \\ >&\frac{1}{T} \Big\{ (\frac{\mu_i + 2\bar{\rho}_{i_{\max}}^k}{2b_i}) l^k - (\frac{1}{b_i} \ln \frac{3\mu_i + 4\bar{\rho}_{i_{\max}}^k}{\mu_i - 2\xi_i^*}) \frac{\mu_i - 2\xi_i^*}{2b_i} \\ &- (T + \frac{1}{b_i} \ln \frac{\mu_i + 2\bar{\rho}_{i_{\max}}^k + \xi_i^*}{\mu_i - 2\xi_i^*}) \frac{\bar{\rho}_{i_{\max}}^k + \xi_i^*}{b_i} \Big\} \ge 0. \end{aligned} \tag{46}$$

Thus, $\left\|\tilde{x}_i^k(t_b^k)\right\|_1 > \bar{e}_i^k(t_b^k)$ holds, and the possibility that the fault l' occurring in system (24) is similar to the trained fault k (for any $k \in \{1, \cdots, N\}/\{l\}$) can be excluded. Consequently, the fault l' will be identified similar to the trained fault l, and the isolation time is obtained as: $t_{iso} = \max\left\{t_b^k, \forall k \in \{1, \cdots, N\}/\{l\}\right\}$. This ends the proof. □

Remark 14. *The isolatability conditions (32)–(33) show that, for different types of faults $\phi_i^{l'}(x_s, u)$ and $\phi_i^k(x_s, u)$ ($l \neq k$), if their dynamic difference $\rho_i^{k,l'}(x_s, u)$ has sufficiently large magnitudes (larger than the bound function $\bar{\rho}_i^k(x_s, u) + 2\xi_i^*$) over some time interval $[t_a^k, t_b^k]$, then, these two types of faults can be effectively isolated. Essentially, these conditions imply that the FI process is achieved by utilizing the known nonlinear function $\bar{\rho}_i^k(x_s, u)$ to separate the faulty dynamics of $\phi_i^{l'}(x_s, u)$ and $\phi_i^k(x_s, u)$.*

VI. Simulation Studies

Consider a typical transport-reaction process in chemical industry, i.e., a long, thin catalytic rod in a reactor, which is borrowed from [6], [19]. The spatio-temporal evolution of the rod temperature is described by the following parabolic PDE:

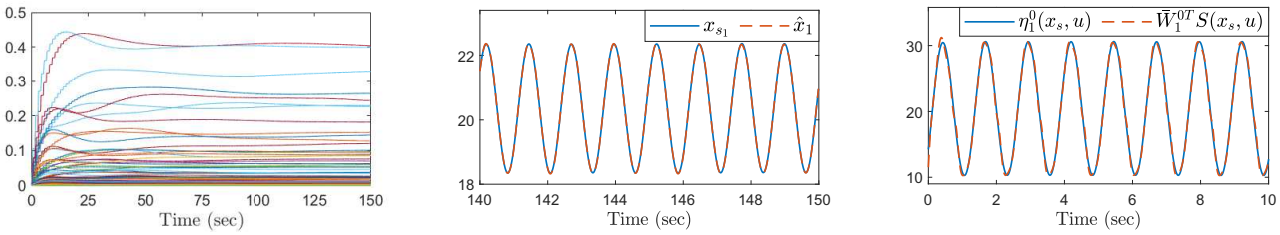
$$\begin{aligned} \frac{\partial x(z,t)}{\partial t} = &\frac{\partial^2 x(z,t)}{\partial z^2} + \beta_T (e^{-\frac{\gamma}{1+x}} - e^{-\gamma}) \\ &+ \beta_u (b(z)u(t) - x(z,t)) + \beta(t - t_0)\phi^k(x,u), \end{aligned} \tag{47}$$

with boundary and initial conditions: $x(0,t) = 0$, $x(\pi,t) = 0$, $x(z,0) = 15\sin(z)$, where $x(z,t)$ denotes the rod temperature, $u(t)$ is the manipulated input, $f(x,u) = \beta_T(e^{-\frac{\gamma}{1+x}} - e^{-\gamma}) + \beta_u(b(z)u - x)$ is an unknown function representing the system uncertainty, and $\phi^k(x,u)$ is the fault function. $\beta_T = 50$ is a heat of action, $\gamma = 4$ is an activation energy, $\beta_u = 2$ is a heat transfer coefficient, $b(z) = 1.5\sin(z) + 1.8\sin(2z) + 2\sin(3z)$ is the actuator distribution function, $t_0 = 30s$ is the fault occurrence time. For simulation purpose, the system input is set as $u(t) = 1.1 + 2\sin(5t) - 2\cos(5t)$. Three types of faults are considered. (i) Fault 1: an actuator fault with a faulty actuator distribution function: $b'(z) = 1.8\sin(z) + 1.8\sin(2z) + 2\sin(3z)$, leading to the associated fault function $\phi^1(x,u) = \tilde{b}(z)\beta_u u$ with $\tilde{b}(z) = b(z) - b'(z) = -0.3\sin(z)$. (ii) Fault 2: a state fault with fault function $\phi^2(x,u) = \tilde{h}(z)x$, where $\tilde{h}(z) = h(z-1) - h(z-1.3)$ and $h(\cdot)$ is a heaviside function. (iii) Fault 3: a component fault with a faulty system parameter: $\beta_T' = 48$, and the associated fault function is thus $\phi^3(x,u) = \tilde{\beta}_T(e^{-\frac{\gamma}{1+x}} - e^{-\gamma})$ with $\tilde{\beta}_T = \beta_T - \beta_T' = 2$.

For the PDE system (47), we first derive its approximate ODE model. Specifically, the eigenvalue problem of the spatial differential operator in (47), i.e., $\mathcal{A}x = \frac{\partial^2 x}{\partial z^2}$, $x \in D(\mathcal{A}) := \{x \in \mathcal{H} \,|\, x(0,t) = 0,\ x(\pi,t) = 0\}$, can be solved analytically by using the method of [34], resulting in the solution: $\lambda_i = -i^2$, $\varphi_i(z) = \sqrt{\frac{2}{\pi}}\sin(iz)$ with $i = 1, \cdots, \infty$. By choosing the first $m = 3$ number of eigenvalues as dominant ones, we can obtain the following ODE system to describe the dominant dynamics of the PDE system (47):

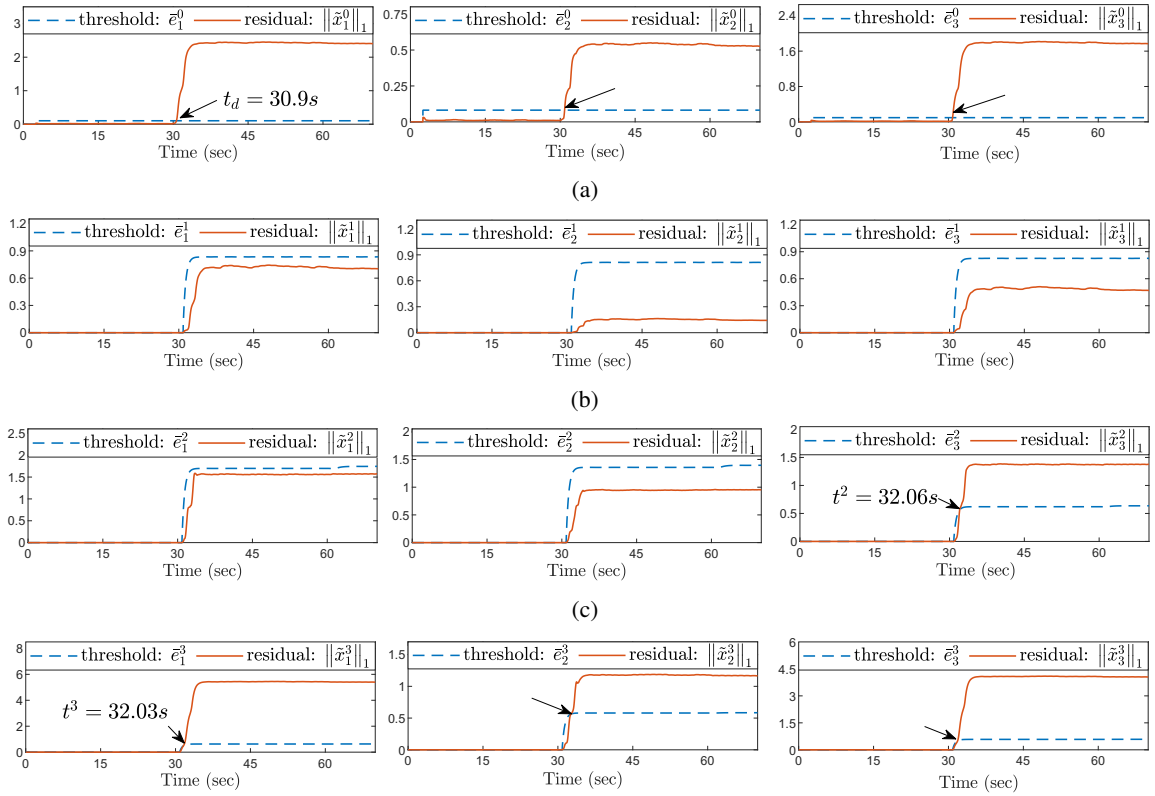
$$\dot{x}_{s_i} = \lambda_i x_{s_i} + f_{s_i}(x_s, u) + \beta(t - t_0)\phi_{s_i}^k(x_s, u), \tag{48}$$

where $i = 1,2,3$, $k = 1,2,3$, $x_{s_i}(t) = \int_0^\pi x(z,t)\varphi_i(z)\,dz$, $f_{s_i}(x_s,u) = -\beta_u x_{s_i}(t) + \int_0^\pi (\beta_T(e^{-\frac{\gamma}{1+\sum_{i=1}^3 x_{s_i}(t)\varphi_i(z)}} - e^{-\gamma}) + \beta_u b(z)u(t))\varphi_i(z)\,dz$, $\phi_{s_i}^1(x_s,u) = \beta_u u(t)\int_0^\pi \tilde{b}(z)\varphi_i(z)\,dz$, $\phi_{s_i}^2(x_s,u) = \int_0^\pi x(z,t)\tilde{h}(z)\varphi_i(z)\,dz$, and $\phi_{s_i}^3(x_s,u) = \tilde{\beta}_T \int_0^\pi (e^{-\frac{\gamma}{1+\sum_{i=1}^3 x_{s_i}(t)\varphi_i(z)}} - e^{-\gamma})\varphi_i(z)\,dz$. Note that the model (48) cannot be directly used for the design of FDI scheme, due to the existence of uncertain functions $f_{s_i}(x_s,u)$ and $\phi_{s_i}^k(x_s,u)$. The state signals x_{s_i}, which are needed for



(a) (b) (c)

Fig. 1: Identification of function $\eta_1^0(x_s,u)$ with identifier (9). (a) Convergence of NN weight $\hat{W}_1^0$; (b) estimation performance of x_{s_1} by $\hat{x}_1$; and (c) function approximation performance of $\eta_1^0(x_s,u)$ by $\bar{W}_1^{0\top}S(x_s,u)$.



(a) (b) (c) (d)

Fig. 2: FDI performance when fault $1'$ occurs at time $t_0 = 30s$ in system (47): (a) FD residuals and thresholds; (b) 1-st FI residuals and thresholds; (c) 2-nd FI residuals and thresholds; and (d) 3-rd FI residuals and thresholds.

the subsequent implementation, will be obtained based on the measurement from the original PDE system (47) via $x_{s_i}(t) = \int_0^\pi x(z,t)\varphi_i(z)\,dz$, as discussed in Remark 2.

With the system signals (x_s, u), we can implement the identification process for the uncertain dynamics $\eta_i^k(x_s,u) = f_{s_i}(x_s,u) + \phi_{s_i}^k(x_s,u)$ $(\forall i = 1,2,3)$ of system (48) under all normal and faulty modes with $k = 0,1,2,3$. Specifically, according to (9), the RBF network $\hat{W}_i^{k\top}S(x_s,u)$ is constructed in a regular lattice, with nodes $N_n = 14\times 9\times 8\times 13$, the center evenly spaced on $[17.5,24]\times[-1,3]\times[0,3.5]\times[-2,4]$ and the widths $\nu_i = 0.5$. The design parameters in (9) are $a_i = 4$, $\Gamma_i = 0.35$ and $\sigma_i = 0.001$ $(\forall i = 1,2,3)$. The initial conditions are set as $\hat{W}_i^k(0) = 0$ and $\hat{x}(0) = x_s(0)$. Consider the system (48) operating in normal mode (with $k = 0$, $\phi_{s_i}^0(x_s,u) \equiv 0$), with identifier (9), the identification performance for the dynamics of the 1-st state subsystem of (48) is shown in Fig. 1. Particularly, Fig. 1a shows the convergence of NN weight $\hat{W}_1^0$. Fig. 1b shows the accurate tracking performance of $\hat{x}_1$ over the system state x_{s_1}. Based on the identification result, a constant model $\bar{W}_1^{0\top}S(x_s,u)$ is obtained by $\bar{W}_1^0 = \frac{1}{10}\int_{140}^{150}\hat{W}_1^0(\tau)\,d\tau$, which can achieve accurate approximation of the associated unknown function $\eta_1^0(x_s,u)$, as illustrated in Fig. 1c. Then, following a similar procedure established as above, simulation results for the cases of faulty modes $k = 1,2,3$ can also be obtained, which are similar to those in Fig. 1 and thus omitted here. Consequently, with the method given in Remark 3, the values of ξ_i^* $(i = 1,2,3)$ that are needed for implementing the

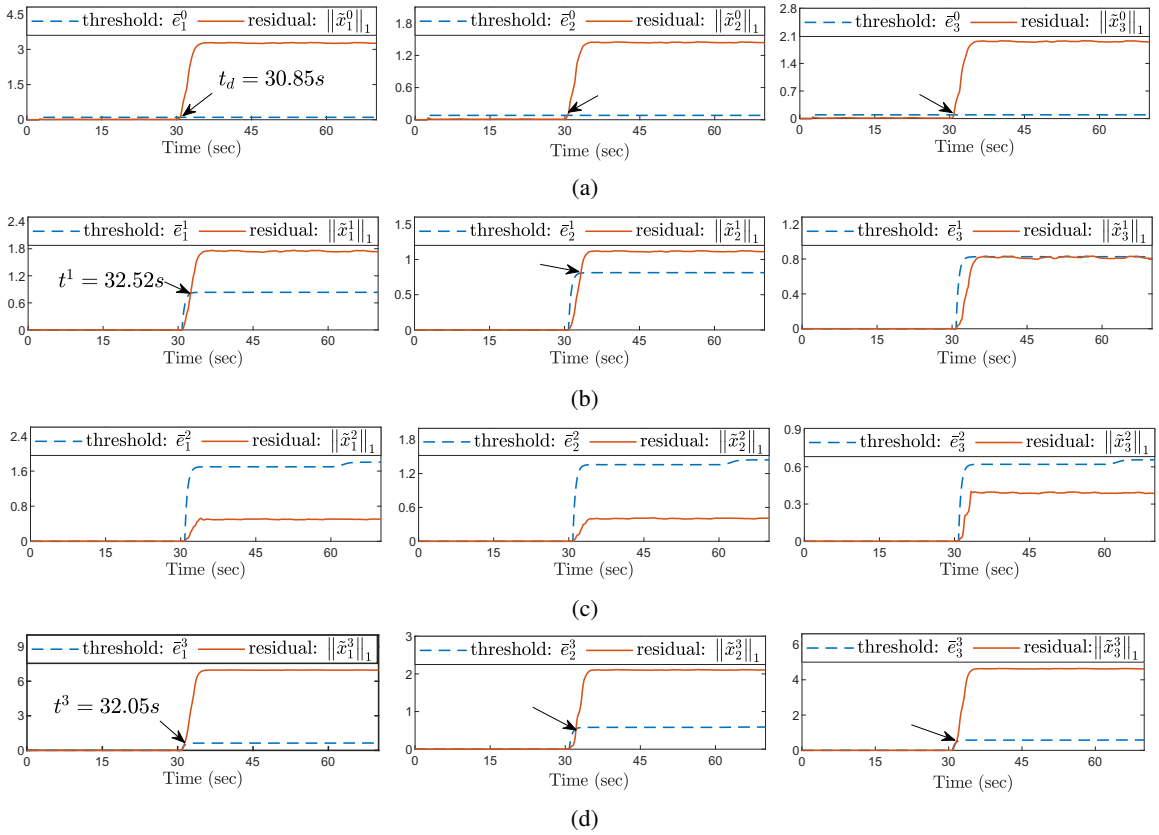


Fig. 3: FDI performance when fault $2'$ occurs at time $t_0 = 30s$ in system (47): (a) FD residuals and thresholds; (b) 1-st FI residuals and thresholds; (c) 2-nd FI residuals and thresholds; and (d) 3-rd FI residuals and thresholds.

subsequent FDI scheme can be obtained as $\xi_1^* = 0.0860$, $\xi_2^* = 0.0430$, and $\xi_3^* = 0.0703$.

Based on the above identification results, we can implement the proposed FDI scheme for system (48). Specifically, the FD estimators (11) are implemented with constant NN models $\bar{W}_i^{0\top} S(x_s, u)$ and parameters $b_i^0 = 2$ ($\forall i = 1, 2, 3$). The FD thresholds (14) are implemented with parameters $\xi_1^* = 0.0860$, $\xi_2^* = 0.0430$, $\xi_3^* = 0.0703$, and $\varrho_i = 0.12$. The parameter of $\mathcal{L}_1$ norm is set as $T = 2.5s$. Similarly, the FI estimators (25) are implemented with constant NN models $\bar{W}_i^{k\top} S(x_s, u)$ and parameters $b_i = 2$ ($\forall i = 1, 2, 3$ and $\forall k = 1, 2, 3$). The FI adaptive thresholds (31) are implemented with given functions $\bar{\rho}_i^k(x_s, u) = \int_0^\pi \bar{\phi}^k(x, u) |\varphi_i(z)| \, dz$, where $\bar{\phi}^1(x, u) = \left|\Delta_{\tilde{b}} \beta_u u\right|$ with $\Delta_{\tilde{b}} = 0.25$, $\bar{\phi}^2(x, u) = \left|\Delta_{\tilde{h}} x\right|$ with $\Delta_{\tilde{h}} = h(z-1) - h(z-1.3)$, and $\bar{\phi}^3(x, u) = \left|\Delta_{\tilde{\beta_T}} (e^{-\frac{\gamma}{1+x}} - e^{-\gamma})\right|$ with $\Delta_{\tilde{\beta_T}} = 1$. For testing purpose, we assume three occurring faults to be detected and isolated, including fault $1'$: $\phi^{1'}(x, u) = \tilde{b}'(z) \beta_u u$ with $\tilde{b}'(z) = -0.5 \sin(z)$; fault $2'$: $\phi^{2'}(x, u) = \tilde{h}'(z) x$ with $\tilde{h}'(z) = h(z-1) - h(z-1.2)$; and fault $3'$: $\phi^{3'}(x, u) = \tilde{\beta_T}' (e^{-\frac{\gamma}{1+x}} - e^{-\gamma})$ with $\tilde{\beta_T}' = 49$. These faults satisfy $\left|\phi^k(x, u) - \phi^{k'}(x, u)\right| \leq \bar{\phi}^k(x, u)$ and $\left|\rho_i^{k,k'}(x_s, u)\right| = \left|\int_0^\pi (\phi^{k'}(x, u) - \phi^k(x, u)) \varphi_i(z) dz\right| \leq \int_0^\pi \bar{\phi}^k(x, u) |\varphi_i(z)| \, dz = \bar{\rho}_i^k(x_s, u)$ for all $k = 1, 2, 3$ and $i = 1, 2, 3$, which verifies Assumption 3.

In the testing phase, consider the fault $1'$ occurring in system (47) at time $t_0 = 30s$, the associated FDI simulation results are displayed in Fig. 2. We first observe the FD performance in Fig. 2a. It is shown that once the fault $1'$ occurs at time $t_0 = 30s$, all FD residuals $\left\|\tilde{x}_i^0\right\|_1$ ($i = 1, 2, 3$) increase and become larger than the associated thresholds $\bar{e}_i^0$ at time $t_d = 30.9s$, indicating that the occurring fault $1'$ can be detected at time $t_d = 30.9s$. Once fault $1'$ is detected, the FI system consisting of FI estimators (25) and FI adaptive thresholds (31) is activated, and the FI performance can be seen in Figs. 2b-2d. For the performance of the matched/similar FI estimator (i.e., the 1-st FI estimator) as shown in Fig. 2b, it is seen that all the residual signals $\left\|\tilde{x}_i^1(t)\right\|_1$ ($i = 1, 2, 3$) remain smaller than the associated threshold $\bar{e}_i^1(t)$ for all time $t > t_d = 30.9s$. For those estimators with unmatched/unsimilar faults (i.e., the 2-nd and the 3-rd FI estimators), the associated performance is presented in Figs. 2c, 2d, respectively. It is shown that the 2-nd FI residual $\left\|\tilde{x}_i^2(t)\right\|_1$ (with $i = 3$) becomes larger than the threshold $\bar{e}_i^2(t)$ at time $t^2 = 32.06s$ (see Fig. 2c); and all the 3-rd FI residuals $\left\|\tilde{x}_i^3(t)\right\|_1$ (with all $i = 1, 2, 3$) become larger than the respective thresholds $\bar{e}_i^3(t)$ at time $t^3 = 32.03s$ (see Fig. 2d). Thus, it can be deduced that the occurring fault $1'$ is similar to the fault 1, and the isolation time can be obtained at: $t_{iso} = \max\{t^2, t^3\} = 32.06s$. Next, we further consider the cases when faults $2'$ and $3'$ are occurring respectively in system (47) at time $t_0 = 30s$, and the associated FDI performances are illustrated in Figs. 3 and 4, respectively. It is seen that the occurring fault $2'$ is detected at time $t_d = 30.85s$ and isolated

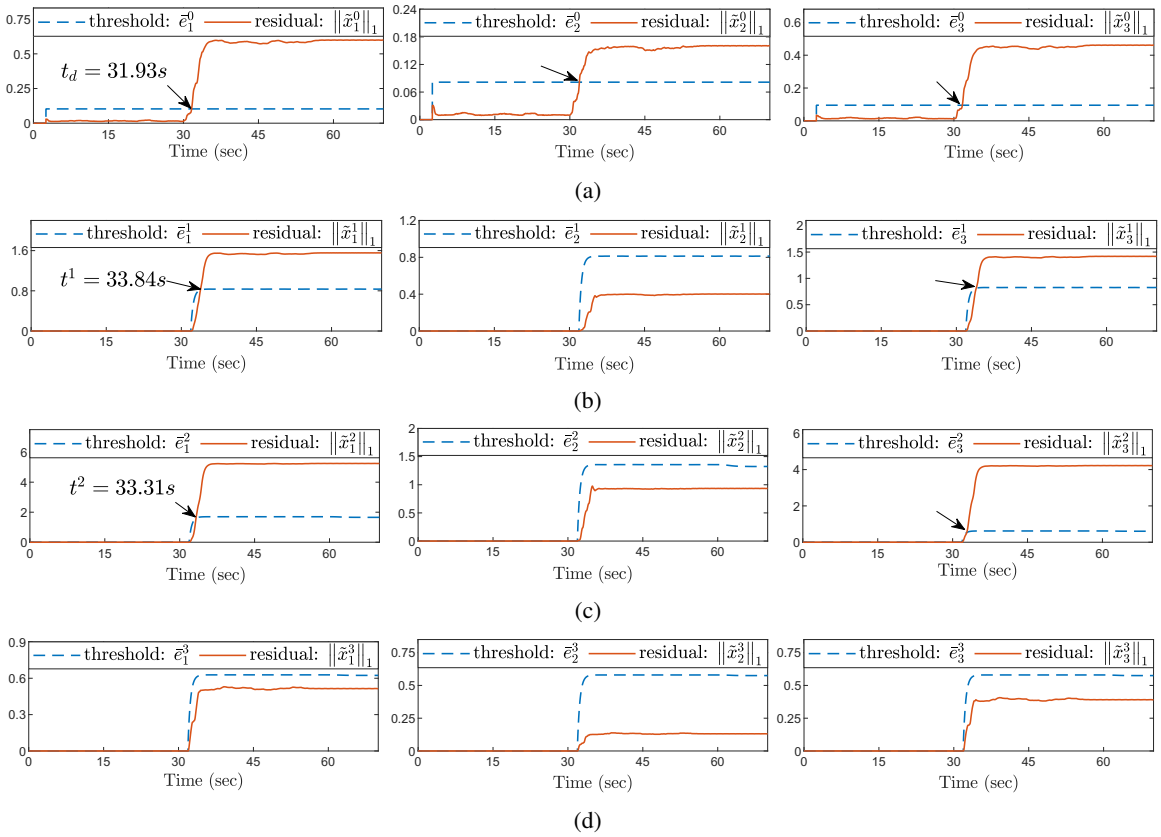
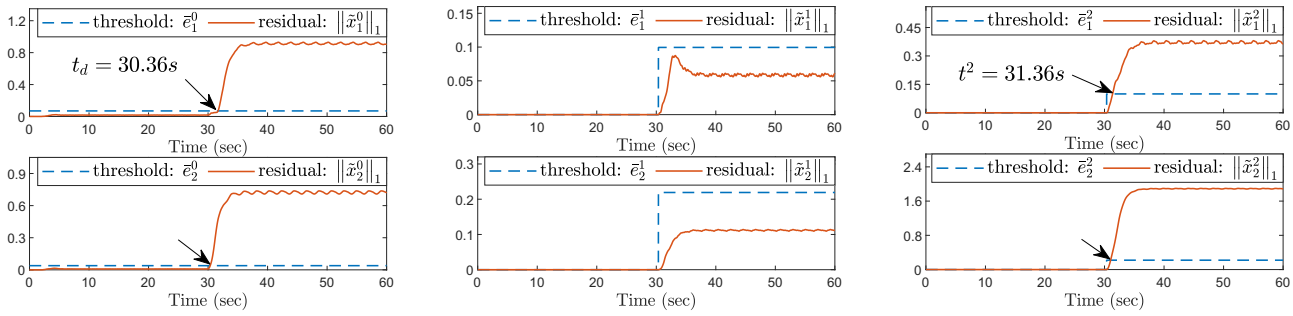


Fig. 4: FDI performance when fault $3'$ occurs at time $t_0 = 30s$ in system (47): (a) FD residuals and thresholds; (b) 1-st FI residuals and thresholds; (c) 2-nd FI residuals and thresholds; and (d) 3-rd FI residuals and thresholds.

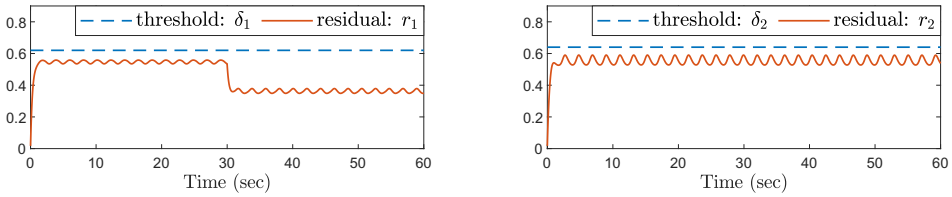
at time $t_{iso} = 32.52s$; while the occurring fault $3'$ is detected at time $t_d = 31.93s$ and isolated at time $t_{iso} = 33.84s$. These simulation results demonstrate feasibility and effectiveness of our proposed FDI scheme.

To further justify the advantage of our FI scheme in dealing with the system uncertainty for accurate isolation, we compare the performance of our scheme with the existing method in [6]. A PDE system in the form of (47) is considered, where the system structures/parameters keep unchanged except $b(z) = [h(z) - h(z-\pi/2), h(z-\pi/2) - h(z-\pi)]$ and $u(t) = [1.1+6\sin(3t); 1.1-6\cos(3t)]$. This system model is assumed partially-unknown, i.e., it consists of (known) nominal component $\frac{\partial^2 x}{\partial z^2} + f(x) + \beta_u b(z)u = \frac{\partial^2 x}{\partial z^2} + 0.9\beta_T(e^{-\frac{\gamma}{1+x}} - e^{-\gamma}) + \beta_u(b(z)u - x)$ and uncertain/unknown component $N(x) = 0.1\beta_T(e^{-\frac{\gamma}{1+x}} - e^{-\gamma})$. Two types of actuator faults at different locations are considered, i.e., $\phi^k(x,u) = \beta_u b(z) f_a^k(x)$ $(k = 1, 2)$ with $f_a^1(x) = [-0.05x(\frac{\pi}{2}, t); 0]$ denoting fault 1 and $f_a^2(x) = [0; -0.05x(\frac{\pi}{2}, t)]$ denoting fault 2. The approximate ODE model of this system is derived with order $m = 2$. For the FI scheme of [6], a bank of FI filters (generating FI residuals $r_i(t)$, $i = 1, 2$) are constructed according to [6, Eq. (27)] and the corresponding FI thresholds are given as $\delta_1 = 0.62$ and $\delta_2 = 0.64$, which are determined based on the upper bound of system uncertainty $N(x)$, according to [6, Remark 19]. For our scheme, the implementation process follows a similar line established as above, in which the RBF network is constructed with nodes $N_n = 13 \times 9 \times 15 \times 15$, the center evenly spaced on $[14, 26] \times [-4, 4] \times [-6, 8] \times [-6, 8]$ and the widths $\nu_i = 1$; and the associated parameters are given as $b_i^0 = b_i = 1$, $\varrho_i = 0.02$ $(i = 1, 2)$, $\xi_1^* = 0.0495$, $\xi_2^* = 0.0191$, $\bar{\rho}_1^1 = 0.05$, $\bar{\rho}_1^2 = 0.2$ and $T = 2s$. For testing purpose, two test faults, i.e., fault $1'$ with $f_a^{1'}(x) = [-0.043x(\frac{\pi}{2}, t); 0]$ and fault $2'$ with $f_a^{2'}(x) = [0; -0.043x(\frac{\pi}{2}, t)]$, are considered occurring at time $t_0 = 30s$. Specifically, considering the case when fault $1'$ occurs, with our scheme, it can be seen in Fig. 5 that the occurring fault $1'$ can be detected at time $t_d = 30.36s$ and be identified similar to fault 1 at time $t_{iso} = t^2 = 31.36s$. With the scheme of [6], it is shown in Fig. 6 that after fault occurrence time $t_0 = 30s$, the matched FI residual $r_1(t)$ do not increase and cross the associated FI threshold δ_1, indicating that isolation for fault $1'$ cannot be achieved. For the case of fault $2'$, similar observations can be seen in Figs. 7-8, where fault $2'$ can be detected at $t_d = 30.38s$ and isolated at $t_{iso} = 32.29s$ with our scheme; but isolation failed with the scheme of [6]. For such results, one important reason lies in that: the FI method of [6] cannot deal with the effect of the system uncertainty $N(x)$ during the FI process, such that the occurring fault dynamics $\phi^k(x,u)$ are hidden within the uncertain dynamics $N(x)$ and cannot be captured for successful isolation; while our method has successfully overcome this issue by achieving accurate identification of system uncertainty $N(x)$. These comparison results are consistent with the discussions in Remarks 6 and 11, demonstrating the advantage of our FI scheme compared



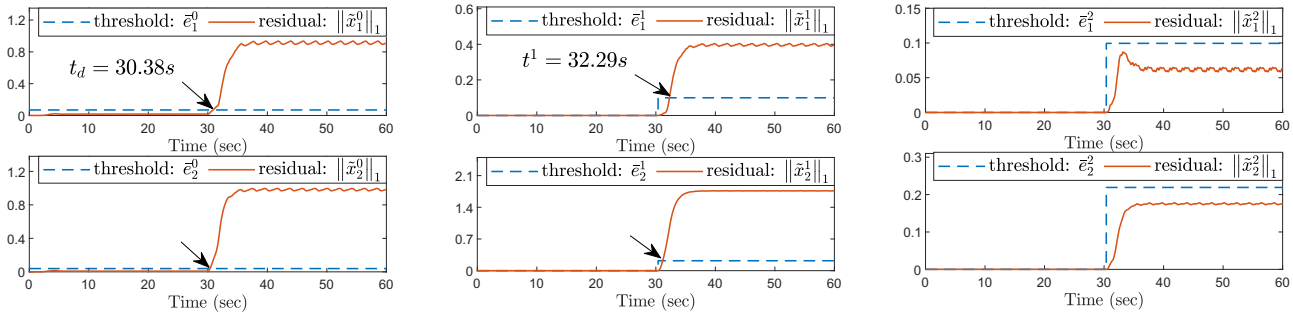
(a) (b) (c)

Fig. 5: FDI performance with the proposed scheme when actuator fault $1'$ occurs at time $t_0 = 30s$ (for comparison study): (a) FD residuals and thresholds; (b) 1-st FI residuals and thresholds; and (c) 2-nd FI residuals and thresholds.



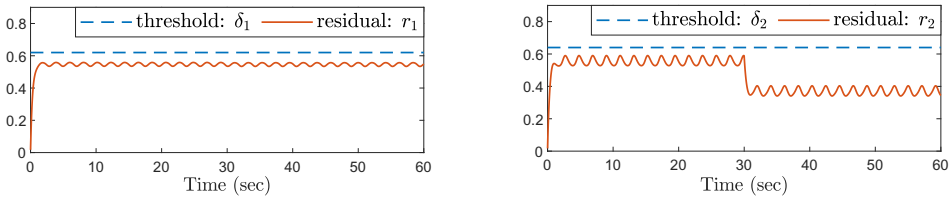
(a) (b)

Fig. 6: FI performance with the method of [6] when actuator fault $1'$ occurs at time $t_0 = 30s$ (for comparison study): (a) 1-st FI residual and threshold; and (b) 2-nd FI residual and threshold.



(a) (b) (c)

Fig. 7: FDI performance with the proposed scheme when actuator fault $2'$ occurs at time $t_0 = 30s$ (for comparison study): (a) FD residuals and thresholds; (b) 1-st FI residuals and thresholds; and (c) 2-nd FI residuals and thresholds.



(a) (b)

Fig. 8: FI performance with the method of [6] when actuator fault $2'$ occurs at time $t_0 = 30s$ (for comparison study): (a) 1-st FI residual and threshold; and (b) 2-nd FI residual and threshold.

to that of [6].

Remark 15. *Comparison study for the proposed FD scheme has been performed in our preliminary work [14], which is thus not repeated here.*

VII. Conclusions

In this paper, we have proposed a novel FDI scheme for a class of uncertain nonlinear parabolic PDE systems. The design was based on an approximate ODE system derived via the Galerkin method, which is used to capture the dominant dynamics of the original PDE system. Specifically, based on the ODE system, a DL-based adaptive dynamics learning approach was first developed to achieve locally-accurate identification of the system uncertain dynamics under normal and all faulty modes. The learned knowledge was obtained and stored in constant RBF NN models. Then, a bank of FDI estimators can be designed with these models. In particular, the FD estimators are used to detect the occurrence of a fault; while the FI estimators, which will be activated once the fault is detected, are used to identify the type of occurring fault. The thresholds associated with these estimators were further designed for real-time decision making. The associated analysis on FDI performance, i.e., fault detectability and isolatability conditions, has also been provided. Simulation studies have been conducted to verify the effectiveness and advantage of the proposed methodologies.

References

[1] Y. Song, X. He, Z. Liu, W. He, C. Sun and F.Y. Wang, "Parallel control of distributed parameter systems," *IEEE Transactions on Cybernetics*, vol. 48, no. 12, p. 3291–3301, 2018.

[2] G. Montaseri and M. J. Yazdanpanah, "Predictive control of uncertain nonlinear parabolic pde systems using a galerkin/neural-network-based model," *Communications in Nonlinear Science and Numerical Simulation*, vol. 17, no. 1, pp. 388–404, 2012.

[3] K. Xu, B. Fan, H. Yang, L. Hu and W. Shen, "Locally weighted principal component analysis-based multimode modeling for complex distributed parameter systems," *IEEE Transactions on Cybernetics*, 2021, doi: 10.1109/TCYB.2021.3061741.

[4] B. Hernandez-Morales and A. Mitchell, "Review of mathematical models of fluid flow, heat transfer, and mass transfer in electroslag remelting process," *Ironmaking & steelmaking*, vol. 26, no. 6, pp. 423–438, 1999.

[5] S. Bououden, M. Chadli, and H. R. Karimi, "Control of uncertain highly nonlinear biological process based on takagi–sugeno fuzzy models," *Signal Processing*, vol. 108, pp. 195–205, 2015.

[6] N. H. El-Farra and S. Ghantasala, "Actuator fault isolation and reconfiguration in transport-reaction processes," *AIChE Journal*, vol. 53, no. 6, pp. 1518–1537, 2007.

[7] X. Lu, W. Zou, and M. Huang, "A novel spatiotemporal ls-svm method for complex distributed parameter systems with applications to curing thermal process," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 3, pp. 1156–1165, 2016.

[8] E. L. Russell, L. H. Chiang, and R. D. Braatz, *Data-driven methods for fault detection and diagnosis in chemical processes*. London, UK: Springer, 2012.

[9] M. Demetriou and A. Armaou, "Dynamic online nonlinear robust detection and accommodation of incipient component faults for nonlinear dissipative distributed processes," *International Journal of Robust and Nonlinear Control*, vol. 22, no. 1, pp. 3–23, 2012.

[10] J. Cai, H. Ferdowsi, and J. Sarangapani, "Model-based fault detection, estimation, and prediction for a class of linear distributed parameter systems," *Automatica*, vol. 66, pp. 122–131, 2016.

[11] F. Fischer and J. Deutscher, "Flatness-based algebraic fault diagnosis for distributed-parameter systems," *Automatica*, vol. 117, p. 108987, 2020.

[12] Y. Feng, Y. Wang, B.C. Wang, and H.X. Li, "Spatial decomposition-based fault detection framework for parabolic-distributed parameter processes," *IEEE Transactions on Cybernetics*, 2021, doi: 10.1109/TCYB.2021.3049453.

[13] S. Dey, H. E. Perez, and S. J. Moura, "Robust fault detection of a class of uncertain linear parabolic pdes," *Automatica*, vol. 107, pp. 502–510, 2019.

[14] J. Zhang, C. Yuan, W. Zeng, P. Stegagno, and C. Wang, "Fault detection of a class of nonlinear uncertain parabolic pde systems," *IEEE Control Systems Letters*, vol. 5, no. 4, pp. 1459–1464, 2020.

[15] A. Baniamerian and K. Khorasani, "Fault detection and isolation of dissipative parabolic pdes: Finite-dimensional geometric approach," in *2012 American Control Conference (ACC)*. IEEE, 2012, pp. 5894–5899.

[16] J. Cai and S. Jagannathan, "Fault isolation in distributed parameter systems modeled by parabolic partial differential equations," in *2016 American Control Conference (ACC)*. IEEE, 2016, pp. 4356–4361.

[17] S. Ghantasala and N. H. El-Farra, "Robust actuator fault isolation and management in constrained uncertain parabolic pde systems," *Automatica*, vol. 45, no. 10, pp. 2368–2373, 2009.

[18] ——, "Detection, isolation and management of actuator faults in parabolic pdes under uncertainty and constraints," in *2007 46th IEEE Conference on Decision and Control*. IEEE, 2007, pp. 878–884.

[19] H.-N. Wu and H.-X. Li, "A galerkin/neural-network-based design of guaranteed cost control for nonlinear distributed parameter systems," *IEEE transactions on neural networks*, vol. 19, no. 5, pp. 795–807, 2008.

[20] C. Qi and H.-X. Li, "Nonlinear dimension reduction based neural modeling for distributed parameter processes," *Chemical Engineering Science*, vol. 64, no. 19, pp. 4164–4170, 2009.

[21] R. Zhang, J. Tao, R. Lu, and Q. Jin, "Decoupled arx and rbf neural network modeling using pca and ga optimization for nonlinear distributed parameter systems," *IEEE transactions on neural networks and learning systems*, vol. 29, no. 2, pp. 457–469, 2016.

[22] C. Yuan and C. Wang, "Persistency of excitation and performance of deterministic learning," *Systems & Control Letters*, vol. 60, no. 12, pp. 952–959, 2011.

[23] C. Wang and D. J. Hill, *Deterministic learning theory for identification, recognition, and control*. Boca Raton, FL, USA: CRC Press, 2009.

[24] ——, "Deterministic learning and rapid dynamical pattern recognition," *IEEE Transactions on Neural Networks*, vol. 18, no. 3, pp. 617–630, 2007.

[25] T. Chen and C. Wang, "Rapid isolation of small oscillation faults via deterministic learning," *International Journal of Adaptive Control and Signal Processing*, vol. 28, no. 3-5, pp. 366–385, 2014.

[26] T. Chen, C. Wang, and D. J. Hill, "Rapid oscillation fault detection and isolation for distributed systems via deterministic learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 6, pp. 1187–1199, 2013.

[27] J. Zhang, Q. Gao, C. Yuan, W. Zeng, S.-L. Dai, and C. Wang, "Similar fault isolation of discrete-time nonlinear uncertain systems: An adaptive threshold based approach," *IEEE Access*, vol. 8, pp. 80 755–80 770, 2020.

[28] J. Zhang, C. Yuan, P. Stegagno, H. He, and C. Wang, "Small fault detection of discrete-time nonlinear uncertain systems," *IEEE Transactions on Cybernetics*, vol. 51, no. 2, pp. 750–764, 2019.

[29] H.-X. Li and C. Qi, "Modeling of distributed parameter systems for applications—a synthesized review from time–space separation," *Journal of Process Control*, vol. 20, no. 8, pp. 891–901, 2010.

[30] P. Benner, M. Ohlberger, A. Cohen, and K. Willcox, *Model reduction and approximation: theory and algorithms*. Society for Industrial and Applied Mathematics, 2017.

[31] P. D. Christofides and P. Daoutidis, "Finite-dimensional control of parabolic pde systems using approximate inertial manifolds," *Journal of mathematical analysis and applications*, vol. 216, no. 2, pp. 398–420, 1997.

[32] A. Armaou and M. A. Demetriou, "Robust detection and accommodation of incipient component and actuator faults in nonlinear distributed processes," *AIChE journal*, vol. 54, no. 10, pp. 2651–2662, 2008.

[33] M. POEWELL, *The Theory of Radial Basis Function Approximation*. Oxford: Clarendon Press, 1992.

[34] H. Deng, H.X. Li and G. Chen, "Spectral-approximation-based intelligent modeling for distributed thermal processes," *IEEE Transactions on Control Systems Technology*, vol. 13, no. 5, p. 686–700, 2005.