Mechanical, Industrial & Systems Engineering
Faculty Publications

Mechanical, Industrial & Systems Engineering

4-28-2020

# Similar Fault Isolation of Discrete-Time Nonlinear Uncertain Systems: An Adaptive Threshold Based Approach

Jingting Zhang

Qingbin Gao

Chengzhi Yuan

Wei Zeng

Shi-Lu Dai

*See next page for additional authors*

## Authors

Jingting Zhang, Qingbin Gao, Chengzhi Yuan, Wei Zeng, Shi-Lu Dai, and Cong Wang

# Similar Fault Isolation of Discrete-Time Nonlinear Uncertain Systems: An Adaptive Threshold Based Approach

**JINGTING ZHANG[1], QINGBIN GAO[2], (Member, IEEE), CHENGZHI YUAN [1], (Member, IEEE), WEI ZENG[3], SHI-LU DAI [4], (Member, IEEE), AND CONG WANG[5], (Member, IEEE)**

[1]Department of Mechanical, Industrial and Systems Engineering, The University of Rhode Island, Kingston, RI 02881, USA
[2]School of Mechanical Engineering and Automation, Harbin Institute of Technology, Shenzhen 518055, China
[3]School of Mechanical and Electrical Engineering, Longyan University, Longyan 364012, China
[4]School of Automation Science and Engineering, South China University of Technology, Guangzhou 510641, China
[5]School of Control Science and Engineering, Shandong University, Jinan 250061, China

Corresponding author: Chengzhi Yuan (cyuan@uri.edu)

**ABSTRACT** In this paper, a new concept of ''similar fault'' is introduced to the field of fault isolation (FI) of discrete-time nonlinear uncertain systems, which defines a new and important class of faults that have small mutual differences in fault magnitude and fault-induced system trajectories. Effective isolation of such similar faults is rather challenging as their small mutual differences could be easily concealed by other system uncertainties (e.g., modeling uncertainty/disturbances). To this end, a novel similar fault isolation (sFI) scheme is proposed based on an adaptive threshold mechanism. Specifically, an adaptive dynamics learning approach based on the deterministic learning theory is first introduced to locally accurately learn/identify the uncertain system dynamics under each faulty mode using radial basis function neural networks (RBF NNs). Based on this, a bank of sFI estimators are then developed using a novel mechanism of absolute measurement of fault dynamics differences. The resulting residual signals can be used to effectively capture the small mutual differences of similar faults and distinguish them from other system uncertainties. Finally, an adaptive threshold is designed for real-time sFI decision making. One important feature of the proposed sFI scheme is that: it is capable of not only isolating similar faults that belong to a pre-defined fault set (used in the training/learning process), but also identifying new faults that do not match any pre-defined faults. Rigorous analysis on isolatability conditions and isolation time is conducted to characterize the performance of the proposed sFI scheme. Simulation results on a practical application example of a single-link flexible joint robot arm are used to show the effectiveness and advantages of the proposed scheme over existing approaches.

**INDEX TERMS** Adaptive dynamics learning, deterministic learning, discrete-time systems, neural networks, nonlinear uncertain systems, similar fault isolation.

## I. INTRODUCTION

Fault isolation (FI) is an important and challenging problem that has been extensively investigated in the systems and control community [1]–[3]. It constitutes one of the critical steps in many complicated control system designs, e.g., fault-tolerant control [4], [5] and condition-based maintenance [6], [7]. The past decades have witnessed considerable research

The associate editor coordinating the review of this manuscript and approving it for publication was Haiquan Zhao .

interests in the field, leading to numerous interesting results on both analysis and design of various FI algorithms (see, e.g., [8]–[19]). In this paper, we will introduce a new concept of ''similar faults'' to the FI field, which defines an important class of faults that have small mutual differences in magnitude and associated system trajectories. The associated similar fault isolation (sFI) problem is of great importance for many modern engineering systems (e.g., [20]–[22]). A notable example is the high precision robotics [23], [24], which is often used for complex assembly or positioning
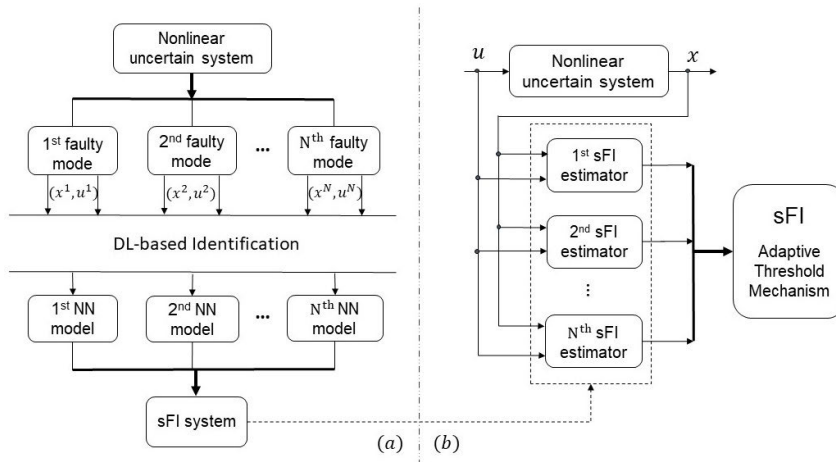
**FIGURE 1.** Architecture of the proposed sFI scheme ($u$ and $x$ represent the system inputs and states, respectively): (a) fault dynamics identification phase; (b) fault isolation phase.

tasks to achieve high-precision motion. For such engineering applications, to meet stringent safety and precision requirements, a reliable sFI scheme is necessary. However, different from the control problem for which system faults could be effectively handled by advanced nonlinear control methods (e.g., [4], [5], [34]), the sFI problem is rather challenging, especially when the system is subject to various system uncertainties such as modeling uncertainty or disturbances. The technical difficulty lies in that for similar faults, their mutual differences are relatively small and could be easily concealed by other system uncertainties. This makes many existing FI schemes not applicable to dealing with the sFI problem. For example, in [15]–[17], since the effects of system uncertainties on measuring fault differences have not been properly dealt with, an isolatable fault is required to be sufficiently different from their mismatched faults, and generally the magnitude of fault differences is required to be larger than that of the system uncertainty, which may not be always feasible under the sFI context.

In this paper, we focus on addressing the sFI problem for discrete-time nonlinear uncertain systems by proposing a novel adaptive threshold based sFI scheme. Specifically, we aim at achieving accurate isolation for various types of similar faults by distinguishing the occurring fault from system uncertainties. As shown in Fig. 1, the proposed sFI scheme consists of two phases, i.e., the fault dynamics identification phase, and fault isolation phase. The identification phase seeks to overcome the system uncertainty issue as discussed above. More specific, considering that the system operates under different faulty modes, the uncertain system dynamics will be accurately identified by using a deterministic learning (DL) based adaptive dynamics learning approach from [25]. The learned knowledge can be stored as a bank of constant radial basis function neural network (RBF NN) models. In the fault isolation phase, with the obtained RBF NN models, a bank of sFI estimators are designed through a novel mechanism of absolute measurement of fault dynamics

difference. Their generated residuals are able to capture the fault differences (between the occurring fault and each trained fault), and distinguish such differences from other system uncertainties for sFI decision making. An adaptive threshold is designed based on these residuals, which can be used to pre-specify the required matching level between an isolatable fault and its matched fault in the training set. The real-time sFI decision making is based on an adaptive threshold mechanism, i.e., the occurring fault is identified similar to a unique trained fault when the related residual becomes smaller than the threshold. Through rigorous analysis, it is demonstrated that with the proposed sFI scheme, different types of similar faults that have been trained can be accurately isolated. In addition, for those faults that do not belong to the training fault set, the proposed sFI scheme is still capable of identifying them as new faults.

It should be noted that this research work is developed by extending our previous work of small fault detection (sFD) in [25]. Different from the sFD problem, the sFI problem considered in this paper is more challenging. The technical difficulty lies in how to effectively evaluate the mismatching/matching level between the occurring fault and each trained fault, especially when considering similar faults. Specifically, since the mutual differences among similar faults are relatively small, to accurately evaluate the fault matching level, it is necessary to study the functional structures of the fault differences in terms of not only the magnitude but also the sign. Most of the existing FI approaches only consider the magnitude factor. For example, in [14], [15], FI is achieved by monitoring the accumulated effect of fault differences. It requires that there is no change of sign and the magnitude of fault differences should be sufficiently large for a sufficiently long time. In the case when the fault differences have frequently-changing signs, their accumulated effect could be offset and approach zero, leading to possible FI misjudgment. In our scheme, this issue is addressed based on a novel mechanism of absolute measurement of

fault dynamics differences. Furthermore, for sFI, a class of promising approaches [18], [19], [26] are developed based on smallest residual principle (SRP), which, however, cannot guarantee the desired fault matching level, leading to poor isolation reliability. For example, with the methods of [18], [19], [26], any kind of faults could be isolated (i.e., identified similar to a trained fault), even for those new faults that do not belong to the pre-defined set of trained faults. This deficiency has been overcome in this paper with our method by using an adaptive threshold mechanism instead of SRP.

The main contributions of this research work are summarized as follows:

1) The sFI problem of discrete-time nonlinear uncertain systems is addressed, where the considered similar faults are allowed to have relatively small mutual differences that could be easily concealed by other system uncertainties.

2) An adaptive threshold mechanism based sFI approach is proposed by developing a bank of sFI estimators through a novel mechanism of absolute measurement of fault dynamics differences, which are capable of capturing the small fault differences and distinguishing them from system uncertainties, and are also able to overcome the aforementioned sign issue suffered by existing FI approaches.

3) Rigorous performance analysis is conducted to derive the fault isolatability conditions and the fault isolation time under the proposed sFI scheme.

The rest of this paper is organized as follows. In Section 2, the problem and some preliminary results are provided, including the adaptive dynamics learning approach for accurate identification of uncertain system dynamics. Section 3 presents the proposed sFI scheme and rigorous performance analysis. Simulation results are shown in Section 4 to verify the effectiveness and advantages of the proposed sFI scheme over existing approaches. Conclusions are made in Section 5.

*Notation:* Throughout this paper, $\mathbb{R}$, $\mathbb{R}_+$ and $\mathbb{Z}_+$ 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; $I_n$ denotes the $n \times n$ identity matrix; the open ball $B_r = \{x \in \mathbb{R}^n : ||x|| < r\}$ with $r$ being an arbitrary positive constant; $|\cdot|$ is the absolute value of a real number; $||\cdot||$ is the 2-norm of a vector or a matrix, i.e. $||x|| = (x^T x)^{\frac{1}{2}}$; $||\cdot||_1$ is the $L_1$-norm of a vector or a matrix, i.e. $||x(k)||_1 = \frac{1}{K} \sum_{h=k-K}^{k-1} |x(h)|$ ($k \geq K > 1$); $\lceil a \rceil$ denotes the least integer greater than or equal to a real number $a$.

## II. PROBLEM FORMULATION AND PRELIMINARIES
### A. PROBLEM FORMULATION
Consider the following discrete-time nonlinear uncertain system:

$$x(k+1) = f(x(k), u(k)) + v(x(k), u(k))$$
$$+ \beta(k - k_0)\phi^s(x(k), u(k)), \quad (1)$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$ are the system state and input, respectively, $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ denotes the nominal healthy system dynamics, $v : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ denotes the modeling system uncertainty, $\phi^s : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ represents the functional structure of the deviation in system dynamics due to the fault $s \in \{1, \cdots, N\}$, $k_0$ is an unknown fault occurrence time and $\beta(k - k_0)$ characterizes a fault time profile, i.e., $\beta(k - k_0) = 0$ for $k < k_0$, meaning that no fault occurs in system (1), and $\beta(k - k_0) = 1$ for $k \geq k_0$, meaning that the fault $s$ occurs in (1).

*Remark 1: In this paper, we assume that the system signals $(x, u)$ in (1) are measurable in real-time for fault isolation. For those cases when full system states are not available (i.e., only partial states or output measurements are available), one can resort to state observer techniques (e.g., [35], [36]), which however is out of the scope of this paper and will be investigated in our future work.*

Without losing any generality, for fault isolation purpose, we consider the case when system (1) operates in the $s$-th faulty mode, i.e., $k \geq k_0$. Note that the fault occurrence time $k_0$ does not need to be pre-known, that is, fault detection does not need to be achived in advance. Motivated by [18], [19], we characterize the so-called "similar faults" $\phi^s(x, u)$ occurring in (1) as: 1) the difference between the fault $s$ and each other fault $\bar{s}$ ($\bar{s} \in \{1, \cdots, N\}/\{s\}$), denoted by $\phi^s(x, u) - \phi^{\bar{s}}(x, u)$, is small, which could be hidden within the modeling uncertainty, that is, the magnitude of each $\phi^s(x, u) - \phi^{\bar{s}}(x, u)$ could be smaller than that of the system uncertain dynamics $v(x, u)$, and 2) the $s$-th faulty system trajectory (denoted by $(x^s, u^s)$) is close to each $\bar{s}$-th faulty system trajectory (denoted by $(x^{\bar{s}}, u^{\bar{s}})$), i.e.,

$$dist\left((x^s, u^s), (x^{\bar{s}}, u^{\bar{s}})\right)$$
$$:= \max\{\min ||(x^s, u^s), (x^{\bar{s}}, u^{\bar{s}})||\} < d_\zeta, \quad (2)$$

where $d_\zeta$ is the constant number satisfying $0 < d_\zeta < d_\zeta^*$ with $d_\zeta^*$ being the size of the NN approximation region to be specified later.

### B. RBF NNs
Radial Basis Function Neural Network (RBF NN) is a neural network architecture that can solve any function approximation problem [27]. Specifically, the RBF networks can be described by $f_{nn}(Z) = \sum_{i=1}^{N_n} w_i s_i(Z) = W^T S(Z)$, where $Z \in \Omega_Z \subset \mathbb{R}^q$ is the input vector, $W = [w_1, \cdots, w_{N_n}]^T \in \mathbb{R}^{N_n}$ is the weight vector, $N_n$ is the NN node number, and $S(Z) = [s_1(||Z - \varsigma_1||), \cdots, s_{N_n}(||Z - \varsigma_{N_n}||)]^T$, with $s_i(\cdot)$ being a radial basis function, and $\varsigma_i$ ($i = 1, 2, \cdots, N_n$) being distinct points in state space. The Gaussian function $s_i(||Z - \varsigma_i||) = \exp[\frac{-(Z - \varsigma_i)^T (Z - \varsigma_i)}{\eta_i^2}]$ is one of the most commonly used radial basis functions, where $\varsigma_i = [\varsigma_{i1}, \varsigma_{i2}, \cdots, \varsigma_{iq}]^T$ is the center of the receptive field. and $\eta_i$ 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 easily seen that $S(Z)$ is bounded and there exists a constant $S_M \in \mathbb{R}_+$ such that $||S(Z)|| \leqslant S_M$ [28].

It has been shown in [27], [29] 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 $\epsilon^* > 0$, $f(Z) = W^{*T}S(Z) + \epsilon$, $\forall Z \in \Omega_Z$, where $|\epsilon| < \epsilon^*$ is the ideal approximation error. The ideal weight vector $W^*$ is an "artificial" quantity required for analysis, and is defined as the value of $W$ that minimizes $|\epsilon|$ for all $Z \in \Omega_Z \subset \mathbb{R}^q$, i.e. $W^* \triangleq argmin_{W \in \mathbb{R}^{N_n}} \{ \sup_{Z \in \Omega_Z} |f(Z) - W^T S(Z)| \}$. Moreover, based on the localization property of RBF NNs [28], for any bounded trajectory $Z(k)$ within the compact set $\Omega_Z$, $f(Z)$ can be approximated by using a limited number of neurons located in a local region along the trajectory: $f(Z) = W_\zeta^{*T} S_\zeta(Z) + \epsilon_\zeta$, where $\epsilon_\zeta$ is the approximation error, with $\epsilon_\zeta = O(\epsilon) = O(\epsilon^*)$, $S_\zeta(Z) = [s_{j1}(Z), \cdots, s_{j\zeta}(Z)]^T \in \mathbb{R}^{N_\zeta}$, $W_\zeta^* = [w_{j1}^*, \cdots, w_{j\zeta}^*]^T \in \mathbb{R}^{N_\zeta}$, $N_\zeta < N_n$, and the integers $j_i = j_1, \cdots, j_\zeta$ are defined by $|s_{j_i}(Z_p)| > \theta$ ($\theta > 0$ is a small positive constant) for some $Z_p \in Z(k)$.

It is shown in [28] that for a localized RBF network $W^T S(Z)$ whose centers are placed on a regular lattice, almost any recurrent trajectory[1] $Z(k)$ can lead to the satisfaction of the PE condition of the regressor subvector $S_\zeta(Z)$. This result is formally summarized in the following lemma.

*Lemma 1 ( [30], [31]): Consider any recurrent trajectory $Z(k)$: $\mathbb{Z}_+ \to \mathbb{R}^q$. $Z(k)$ remains in a bounded compact set $\Omega_Z \subset \mathbb{R}^q$, then for RBF network $W^T S(Z)$ with centers placed on a regular lattice (large enough to cover compact set $\Omega_Z$), the regressor subvector $S_\zeta(Z)$ consisting of RBFs with centers located in a small neighborhood of $Z(k)$ is persistently exciting (PE).*

### C. IDENTIFICATION OF FAULT DYNAMICS VIA DETERMINISTIC LEARNING

In our previous work [25], an adaptive dynamics learning approach has been proposed based on DL theory. It can be employed for the identification of the system uncertain dynamics under various faulty modes (i.e., the identification phase as shown in Fig. 1). Specifically, consider the following faulty dynamic systems:

$$x^s(k+1) = f(x^s(k), u^s(k)) + v(x^s(k), u^s(k)) \\ + \phi^s(x^s(k), u^s(k)), \quad (3)$$

where $s = 1, \cdots, N$ is used to stand for the $s$-th faulty mode. Note that the modeling uncertainty $v(x, u)$ and fault function $\phi^s(x, u)$ in (3) cannot be decoupled from each other, we introduce a so-called general fault function $\eta^s(x, u) := v(x, u) + \phi^s(x, u)$, and rewrite the system (3) as:

$$x^s(k+1) = f(x^s(k), u^s(k)) + \eta^s(x^s(k), u^s(k)). \quad (4)$$

For this system, as typically adopted in the DL literature [25], [28], the following assumption is made.

---
[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 [28].

*Assumption 1: The faulty system trajectories $(x^s, u^s)$ $(s = 1, \cdots, N)$ of (4) are all bounded and recurrent.*

Consider the function $\eta^s(x, u) = [\eta_1^s(x, u), \cdots, \eta_n^s(x, u)]^T$ in (4), it is known that there exists an ideal constant RBF NN weight $W^{s*} = [W_1^{s*}, \cdots, W_n^{s*}] \in \mathbb{R}^{N_n \times n}$ (with $N_n$ denoting the number of NN nodes) such that

$$\eta_i^s(x, u) = W_i^{s*T} S(x, u) + \epsilon_{i,0}^s, \quad (5)$$

where $i = 1, \cdots, n$, $s = 1, \cdots, N$, $S(x, u) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^{N_n}$ is a smooth RBF vector and $\epsilon_{i,0}^s$ is the estimation error satisfying $|\epsilon_{i,0}^s| < \epsilon^*$ with $\epsilon^*$ being a positive constant that can be made arbitrarily small given sufficiently large number of neurons. Then, according to [25], we construct the following dynamical identifier:

$$\hat{x}_i(k+1) = a_i(\hat{x}_i(k) - x_i^s(k)) + f_i(x^s(k), u^s(k)) \\ + \hat{W}_i^{sT}(k)S(x^s(k), u^s(k)), \quad (6)$$

where $i = 1, \cdots, n$, $s = 1, \cdots, N$, $\hat{x}_i \in \mathbb{R}$ is the $i$-th state of the identifier, $x_i^s$ is the $i$-th state of the system (4), $0 < a_i < \frac{\sqrt{5}-1}{2}$ is a design parameter, and $\hat{W}_i^s \in \mathbb{R}^{N_n}$ is the estimate of $W_i^{s*}$ which is updated in real time using the following adaptive learning law:

$$\hat{W}_i^s(k+1) = \hat{W}_i^s(k) - c_i \tilde{x}_i(k+1)S(x^s(k), u^s(k)), \quad (7)$$

where $0 < c_i < \frac{1}{S_M^2(2+a_i)}$ is a design constant with $S_M$ being the upper bound of $||S(x^s, u^s)||$, and $\tilde{x}_i := \hat{x}_i - x_i^s$.

The rigorous analysis of the learning performance achieved by the identifier (6)–(7) has been provided in [25]. Specifically, from Lemma 1, the partial PE condition of the regression vector $S(x^s, u^s)$ is guaranteed by the given recurrent trajectory $(x^s, u^s)$ generated from the system (4). Based on this, from [25], the estimation error $\tilde{x}_i$ will converge to a small neighborhood around the origin, and the constructed $\hat{W}_i^{sT} S(x, u)$ will achieve locally-accurate approximation for $\eta_i^s(x^s, u^s)$. In addition, according to [25], the weights $\hat{W}_i^s$ will converge to a small neighborhood of $W_i^{s*}$, and a constant NN weight $\bar{W}_i^s$ can be derived by $\bar{W}_i^s := \frac{1}{K_2} \sum_{k=K_1}^{K_1+K_2-1} \hat{W}_i^s(k)$ with $[K_1, K_1 + K_2 - 1]$ representing a time segment after the transient process. Consequently, a constant model $\bar{W}_i^{sT} S(x^s, u^s)$ can be obtained to represent $\eta_i^s(x^s, u^s)$, i.e.,

$$\eta_i^s(x^s, u^s) = \bar{W}_i^{sT} S(x^s, u^s) + \epsilon_{i,1}^s, \quad (8)$$

with the estimation error $\epsilon_{i,1}^s$ satisfying $|\epsilon_{i,1}^s| \le \varepsilon_i^*$, where $\varepsilon_i^*$ is an arbitrarily small constant which can always be pre-designed and obtained by constructing a sufficiently large number of neurons, as argued in [25].

Furthermore, according to [25], thanks to the generalization ability of constant RBF NN model $\bar{W}_i^{sT} S(x, u)$ in (8), its locally-accurate approximation for $\eta_i^s(x, u)$ can be achieved in a local region

$$\Omega_\zeta^s := \{(x, u) \big| dist((x, u), (x^s, u^s)) < d_\zeta^* \}$$

along the trajectory $(x^s, u^s)$, with $d_\zeta^* > 0$ characterizing the size of the NN approximation region. In such a region,

a desirable accuracy level of the NN approximation (denoted by $\xi_i^*$) will be guaranteed, i.e.,

$$|\eta_i^s(x, u) - \bar{W}_i^{sT} S(x, u)| \leq \xi_i^*, \quad \forall (x, u) \in \Omega_\xi^s.$$

Under the similar fault definition specified in (2), we have that each faulty trajectory $(x^{\bar{s}}, u^{\bar{s}})$ (generated from the $\bar{s}$-th faulty system, $\bar{s} \in \{1, \cdots, N\}/\{s\}$) satisfies $dist((x^s, u^s), (x^{\bar{s}}, u^{\bar{s}})) < d_\zeta^*$, i.e., $(x^{\bar{s}}, u^{\bar{s}}) \in \Omega_\zeta^s$. Thus, we obtain:

$$\eta_i^s(x^{\bar{s}}, u^{\bar{s}}) = \bar{W}_i^{sT} S(x^{\bar{s}}, u^{\bar{s}}) + \epsilon_{i,2}^s, \quad |\epsilon_{i,2}^s| \leq \xi_i^*. \quad (9)$$

In view of (8) and (9), it is demonstrated that no matter which faulty mode the system (1) operates in (i.e., $(x, u) = (x^s, u^s)$ or $(x, u) = (x^{\bar{s}}, u^{\bar{s}})$), the constant RBF NN model $\bar{W}_i^{sT} S(x, u)$ is able to reconstruct the general fault function $\eta_i^s(x, u)$ with desired approximation accuracy. Consequently, the $s$-th faulty system (4) can be represented by:

$$x_i^s(k + 1) = f_i(x(k), u(k)) + \bar{W}_i^{sT} S(x(k), u(k))$$
$$+ \epsilon_i^s, \quad i = 1, \cdots, n, \ s = 1, \cdots, N, \quad (10)$$

where $x_i^s$ represents the $s$-th faulty system state, $(x, u)$ is the trajectory generated in real time from system (1), $\epsilon_i^s$ is the approximation error associated with $\bar{W}_i^{sT} S(x, u)$ satisfying (i) $|\epsilon_i^s| = |\epsilon_{i,1}^s| \leq \varepsilon_i^*$ when the system (1) operates in $s$-th faulty mode, i.e., $(x, u) = (x^s, u^s)$, and (ii) $|\epsilon_i^s| = |\epsilon_{i,2}^s| \leq \xi_i^*$ when the system (1) operates in $\bar{s}$-th faulty mode, i.e., $(x, u) = (x^{\bar{s}}, u^{\bar{s}})$.

## III. MAIN RESULTS

In this section, based on the above learning results, a bank of novel sFI estimators will be designed and a sFI scheme will be developed based on adaptive threshold mechanism. Rigorous analysis of the performance of the proposed sFI scheme will also be provided.

### A. DESIGN OF SIMILAR FAULT ISOLATION ESTIMATOR AND DECISION MAKING SCHEME

Assuming a fault $l'$ (similar to the trained fault $l$, $l \in \{1, \cdots, N\}$) occurs, the monitored system is obtained as:

$$x(k + 1) = f(x(k), u(k)) + v(x(k), u(k))$$
$$+ \beta(k - k_0)\phi^{l'}(x(k), u(k)). \quad (11)$$

Recall that $k_0$ is the fault occurrence time. For this system, by utilizing the constant NN model $\bar{W}_i^{sT} S(x, u)$ ($i = 1, \cdots, n, s = 1, \cdots, N$) obtained from the learning phase, we propose to construct a bank of sFI estimators (i.e., residual systems) embedded with a novel mechanism of absolute measurement of faulty dynamics differences as follows:

$$e_i^s(k) = b_i e_i^s(k - 1) + |f_i(x(k - 1), u(k - 1))$$
$$+ \bar{W}_i^{sT} S(x(k - 1), u(k - 1)) - x_i(k)|, \quad (12)$$

where $i = 1, \cdots, n, s = 1, \cdots, N$; $e_i^s$ is the $i$-th state of the $s$-th residual system with initial conditions set as $e_i^s(0) = 0$; $b_i$ is a design parameter satisfying $0 \leq b_i < 1$; $x_i$ is the $i$-th state of

the monitored system (11). Note that $f_i(x, u) + \bar{W}_i^{sT} S(x, u)$ is able to represent the $s$-th faulty system (4), as verified in (10), the term $|f_i(x(k - 1), u(k - 1)) + \bar{W}_i^{sT} S(x(k - 1), u(k - 1)) - x_i(k)|$ in (12) can thus be used to represent the absolute difference of the dynamics between the $s$-th faulty system (4) and the monitored system (11).

From Eqs. (10) and (11), the residual systems (12) can be rewritten as follows:

$$e_i^s(k) = b_i e_i^s(k - 1) + |\phi_i^s(x(k - 1), u(k - 1))$$
$$- \beta(k - 1 - k_0)\phi_i^{l'}(x(k - 1), u(k - 1)) - \epsilon_i^s|. \quad (13)$$

Introducing a so-called fault mismatch function: $\varrho_i^{s,l'}(x, u) := \phi_i^s(x, u) - \phi_i^{l'}(x, u)$, Eq. (13) is further equivalent to:

$$e_i^s(k) = b_i^{k-k_0} e_i^s(k_0) + \sum_{h=k_0}^{k-1} b_i^{k-1-h} |\varrho_i^{s,l'}(x(h), u(h)) - \epsilon_i^s|. \quad (14)$$

where $\epsilon_i^s$ is the approximation error derived from (10) and can be made small.

*Remark 2: The differences between the occurring fault $l'$ and the trained fault $s$ can be represented by the fault mismatch function $\varrho_i^{s,l'}(x, u)$ ($i = 1, \cdots, n$). The accumulated effects of such fault differences over the time $[k_0, k - 1]$ are characterized by the residual $e_i^s$ in (14), which essentially manifests the matching level between the faults $l'$ and $s$.*

*Remark 3: In the residual system (12), the constant model $\bar{W}_i^{sT} S(x, u)$ ($i = 1, \cdots, n$) is employed to deal with the system uncertainty $v_i(x, u)$ in (11). This enables that the differences between the occurring fault $l'$ and each trained fault $s$ can be captured and distinguished from the system uncertainty $v_i(x, u)$ by the residual $e_i^s$ of (14). Furthermore, the residual system (12) is designed through a novel mechanism of absolute measurement of fault dynamics differences. This further enables that even when the fault difference $\varrho_i^{s,l'}(x, u)$ has frequently-changing signs, its accumulated effects would not be offset and the related residual will not approach zero.*

Based on the residual systems (14), an adaptive threshold will be further developed for real-time sFI decision making. Such an adaptive threshold is a time-varying function used to bound the matched residual signal, i.e., $e_i^l(k)$ ($k \geq k_0$). For this purpose, we first consider an ideal scenario that the fault occurrence time is $k_0 = 0$ and the occurring fault $l'$ completely matches the trained fault $l$, i.e., $\varrho_i^{l,l'}(x, u) \equiv 0$. Consider the $l$-th residual system in the form of (14). Noting that the system (11) operates in the $l$-th faulty mode for all $k \geq 0$, i.e., $(x, u) = (x^l, u^l)$, and $|\epsilon_i^l| = |\epsilon_{i,1}^l| \leq \varepsilon_i^*$ from (10), the residual $e_i^l(k)$ in the form of (14) is obtained as:

$$e_i^l(k) = \sum_{h=0}^{k-1} b_i^{k-1-h} |\epsilon_i^l| \leq \sum_{h=0}^{k-1} b_i^{k-1-h} \varepsilon_i^* = \frac{(1 - b_i^k)\varepsilon_i^*}{1 - b_i}. \quad (15)$$
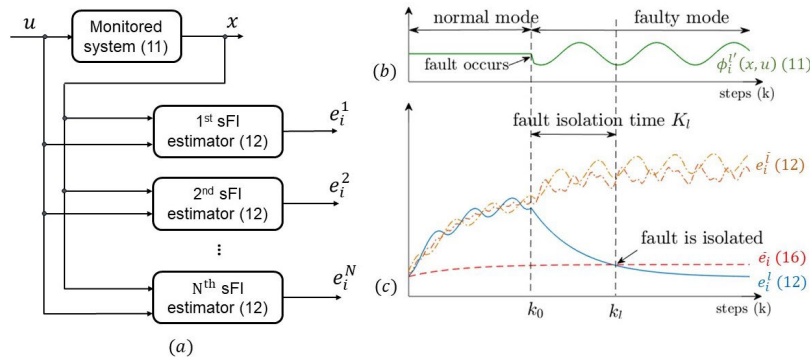
**FIGURE 2.** Schematic of adaptive threshold based sFI scheme: (a) architecture of sFI system; (b) fault function $\phi_i^{l'}(x, u)$ ($i = 1, \cdots, n$); (c) sFI decision making: the matched residual $e_i^l$ ($l \in \{1, \cdots, N\}$), mismatched residuals $e_i^{\bar{l}}$ ($\bar{l} \in \{1, \cdots, N\}/\{l\}$) and adaptive threshold $\bar{e}_i$. Fault $l'$ occurs at time $k_0$ and is isolated at time $k_l$ with $e_i^l(k_l) < \bar{e}_i(k_l)$, the corresponding isolation time $K_l = k_l - k_0$.

Based on (15), we further consider the general case when $k_0 \neq 0$ and $\varrho_i^{l,l'}(x, u) \neq 0$, and design an adaptive threshold as:

$$\bar{e}_i(k) := \frac{(1 - b_i^k)q_i \varepsilon_i^*}{1 - b_i}, \quad i = 1, \cdots, n, \qquad (16)$$

where $q_i > 1$ is an auxiliary design parameter, $b_i$ is the parameter designed in residual systems (12), and $\varepsilon_i^*$ represents the desired accuracy level of the NN approximation, as defined in (10), which can be pre-specified and obtained by setting a sufficiently large number of neurons in the identification phase (an example will be given in the simulation section for illustration).

*Remark 4: In the adaptive threshold (16), the parameter $q_i$ is designed to pre-specify the required matching level between an isolatable fault (fault $l'$) and its matched fault (fault $l$). If the value of $q_i$ is selected small (close to 1), for successful isolation, the fault $l'$ is required to perfectly match the fault $l$ (i.e., $\varrho_i^{l,l'}(x, u) \approx 0$), as verified in (15). Rigorous analysis on the effect of $q_i$ on the isolation performance (in terms of isolation accuracy and isolation time) will be conducted in the next section.*

Based on the residual systems (12) and the threshold (16), an adaptive threshold based sFI scheme can be proposed, as illustrated in Fig. 2. Specifically, after a fault $l'$ occurs (i.e., $k \geq k_0$), the matched residual $e_i^l(k)$ will decrease and become very small such that $e_i^l(k) < \bar{e}_i(k)$ for $k \geq k_l$ (with $k_l$ denoting absolute isolation time), whereas other mismatched residuals $e_i^{\bar{l}}(k)$ ($\bar{l} \in \{1, \cdots, N\}/\{l\}$) will remain large and satisfy $e_i^{\bar{l}}(k) \geq \bar{e}_i(k)$ for all $k \geq k_0$. The idea is formalized as follows:

*sFI Decision Making Scheme:* Compare the residual signals $e_i^s(k)$ ($i = 1, \cdots, n$, $s = 1, \cdots, N$) in (12) with the adaptive threshold $\bar{e}_i(k)$ in (16). For all $i \in \{1, \cdots, n\}$, if there exists a unique $l \in \{1, \cdots, N\}$ and a finite time $k_l > 0$ such that $e_i^l(k_l) < \bar{e}_i(k_l)$, then the occurring fault will be identified similar to fault $l$ at time $k_l$.

*Remark 5: Existing FI methods (e.g., [18], [19] are developed based the SRP, with which any possibly occurring faults, even a new fault that does not match any trained faults, will be identified similar to a trained fault, leading to a possible isolation misjudgment. Advanced over these schemes, our sFI scheme is based on an adaptive threshold mechanism, which can ensure desired isolation accuracy. More specific, only when the difference between the occurring fault $l'$ and the trained fault $l$ is sufficiently small, the related residual $e_i^l$ ($i = 1, \cdots, n$) could become smaller than the threshold $\bar{e}_i$. As for those occasions when no residual becomes smaller than the threshold, the occurring fault will be identified as a new fault.*

*Remark 6: With the proposed sFI scheme based on the adaptive threshold mechanism, similar fault isolation can be achieved without implementing fault detection in advance. This is because when no fault occurs, all residuals will remain larger than the adaptive threshold; when a trained fault recurs, the associated residual will then become smaller than the threshold, leading to successful fault isolation with no need of a priori fault detection.*

### B. FAULT ISOLATABILITY CONDITION AND ISOLATION TIME

In this section, we focus on studying the performance of the proposed sFI scheme in terms of isolatability condition and isolation time. For isolatability, we will examine under what conditions the occurring fault $l'$ could be identified similar to a unique trained fault $l$.

*Theorem 1 (Fault Isolatability): Consider the monitored system (11) and the fault isolation system consisting of (12) and (16). Under the proposed sFI scheme, if the following conditions hold:*

1) *there exists a unique $l \in \{1, \cdots, N\}$ and a finite time $k_f \geq k_0 + \log_{b_i} \frac{(q_i - 1)\varepsilon_i^*}{(q_i - 1)\varepsilon_i^* + (1 - b_i)E_i^l}$ with*

$E_i^l := \max_{k \le k_0}\{e_i^l(k)\}$, *such that for* $\forall i \in \{1, \cdots, n\}$,

$$\sum_{h=k_0}^{k_f-1} b_i^{k_f-1-h}|\varrho_i^{l,l'}(x(h),\, u(h))|$$

$$\le \frac{(1-b_i^{k_f-k_0})(q_i-1)\varepsilon_i^*}{1-b_i} - b_i^{k_f-k_0}E_i^l; \quad (17)$$

2) *for each mismatched fault* $\bar{l}$ ($\bar{l} \in \{1, \cdots, N\}/\{l\}$), *there exist some* $i \in \{1, \cdots, n\}$ *such that*

$$\sum_{h=k_0}^{k-1} b_i^{k-1-h}|\varrho_i^{\bar{l},l'}(x(h),\, u(h))| \ge \frac{q_i\varepsilon_i^* + \xi_i^*}{1-b_i}, \quad \forall k > k_0; \tag{18}$$

*then we have* $e_i^l(k_f) < \bar{e}_i(k_f)$ *and* $e_i^{\bar{l}}(k_f) > \bar{e}_i(k_f)$, *i.e., the occurring fault* $l'$ *is identified similar to the unique trained fault* $l$ *at time* $k_f$.

*Proof:* We first consider the $l$-th residual system in the form of (14). Noting that $k > k_0$, the monitored system (11) operates in $l$-th faulty mode, i.e., $(x, u) = (x^l, u^l)$, $|\epsilon_i^l| = |\epsilon_{i,1}^l| \le \varepsilon_i^*$ from (10), and $e_i^l(k_0) \le E_i^l$, we obtain:

$$e_i^l(k_f)$$

$$= b_i^{k_f-k_0}e_i^l(k_0) + \sum_{h=k_0}^{k_f-1} b_i^{k_f-1-h}|\varrho_i^{l,l'}(x(h),\, u(h)) - \epsilon_i^l|$$

$$\le b_i^{k_f-k_0}e_i^l(k_0) + \sum_{h=k_0}^{k_f-1} b_i^{k_f-1-h}(|\varrho_i^{l,l'}(x(h),\, u(h))| + |\epsilon_i^l|)$$

$$\le b_i^{k_f-k_0}E_i^l + \sum_{h=k_0}^{k_f-1} b_i^{k_f-1-h}|\varrho_i^{l,l'}(x(h),\, u(h))| + \sum_{h=k_0}^{k_f-1} b_i^{k_f-1-h}\varepsilon_i^*$$

$$= b_i^{k_f-k_0}E_i^l + \sum_{h=k_0}^{k_f-1} b_i^{k_f-1-h}|\varrho_i^{l,l'}(x(h),\, u(h))| + \frac{(1-b_i^{k_f-k_0})\varepsilon_i^*}{1-b_i}. \tag{19}$$

Under condition (17), we have: $e_i^l(k_f) \le \frac{(1-b_i^{k_f-k_0})q_i\varepsilon_i^*}{1-b_i} < \frac{(1-b_i^{k_f})q_i\varepsilon_i^*}{1-b_i} = \bar{e}_i(k_f)$. It should be noted that the time instant $k_f$ is required to satisfy $k_f \ge k_0 + \log_{b_i} \frac{(q_i-1)\varepsilon_i^*}{(q_i-1)\varepsilon_i^*+(1-b_i)E_i^l}$, such that the right-hand side of condition (17) is larger than zero, i.e., $\frac{(1-b_i^{k_f-k_0})(q_i-1)\varepsilon_i^*}{1-b_i} - b_i^{k_f-k_0}E_i^l > 0$, guaranteeing the existence of $\sum_{h=k_0}^{k_f-1} b_i^{k_f-1-h}|\varrho_i^{l,l'}(x(h),\, u(h))|$.

Then we consider the $\bar{l}$-th ($\bar{l} \in \{1, \cdots, N\}/\{l\}$) residual system in the form of (14). Recall that $k > k_0$, $(x, u) = (x^l, u^l)$, and $|\epsilon_i^{\bar{l}}| = |\epsilon_{i,2}^{\bar{l}}| \le \xi_i^*$ from (10), we have:

$$e_i^{\bar{l}}(k) = b_i^{k-k_0}e_i^{\bar{l}}(k_0) + \sum_{h=k_0}^{k-1} b_i^{k-1-h}|\varrho_i^{\bar{l},l'}(x(h),\, u(h)) - \epsilon_i^{\bar{l}}|$$

$$\ge \sum_{h=k_0}^{k-1} b_i^{k-1-h}|\varrho_i^{\bar{l},l'}(x(h),\, u(h))| - \sum_{h=k_0}^{k-1} b_i^{k-1-h}|\epsilon_i^{\bar{l}}|$$

$$\ge \sum_{h=k_0}^{k-1} b_i^{k-1-h}|\varrho_i^{\bar{l},l'}(x(h),\, u(h))| - \sum_{h=k_0}^{k-1} b_i^{k-1-h}\xi_i^*$$

$$> \sum_{h=k_0}^{k-1} b_i^{k-1-h}|\varrho_i^{\bar{l},l'}(x(h),\, u(h))| - \frac{\xi_i^*}{1-b_i}. \tag{20}$$

Under condition (18), we obtain $e_i^{\bar{l}}(k) \ge \frac{q_i\varepsilon_i^*}{1-b_i} > \frac{(1-b_i^k)q_i\varepsilon_i^*}{1-b_i} = \bar{e}_i(k)$ for all $k > k_0$, thus, the occurring fault $l'$ will never be identified similar to the mismatched fault $\bar{l}$ under our sFI scheme.

In view of (19) and (20), we have $e_i^l(k_f) < \bar{e}_i(k_f)$ and $e_i^{\bar{l}}(k_f) > \bar{e}_i(k_f)$, thus, the occurring fault $l'$ will be identified similar to a unique trained fault $l$ at time $k_f$. This ends the proof. □

*Remark 7:* Our approach is developed by considering a general case that the occurring fault $l'$ in system (11) does not necessarily have a perfect match to the trained fault $l$ (i.e., $\varrho_i^{l,l'}(x, u) \ne 0$). The isolatability condition (17) is thus proposed to restrict the matching level between the occurring fault $l'$ and the matched fault $l$. Particularly, if we consider a special case that the occurring fault $l'$ is completely identical to the trained fault $l$, i.e., $\varrho_i^{l,l'}(x, u) \equiv 0$ (as typically adopted in the literature [15], [16]), the condition (17) will be automatically satisfied and thus can be omitted.

*Remark 8:* For the isolatability conditions (17)–(18), it should be noted that the bounds given therein can be made small since they are dependent on the parameters $\varepsilon_i^*$ and $\xi_i^*$ (both of which can be made arbitrarily small), and the designed constants $q_i$, $b_i$ (both of which can be selected freely by $q_i > 1$, $1 > b_i \ge 0$). Such small bounds ensure that the condition (18) can be satisfied even when $|\varrho_i^{\bar{l},l'}(x, u)|$ is small.

In the following, we denote the absolute fault isolation time as $k_l$, and the fault isolation time $K_l$ as the time length between the occurrence time $k_0$ and the absolute fault isolation time $k_l$, i.e., $K_l = k_l - k_0$. To further investigate the performance of the proposed sFI scheme, we will derive the upper bound on the isolation time $K_l$ (denoted by $\bar{K}_l$), i.e., sFI will be achieved at most in $\bar{K}_l$ amount of time steps after the fault occurrence time $k_0$.

*Theorem 2 (Isolation Time):* Consider the monitored system (11) and the fault isolation system consisting of (12) and (16). Under the proposed sFI scheme, if the following conditions hold:

1) *for* $\forall i \in \{1, \cdots, n\}$, *there exists a unique* $l \in \{1, \cdots, N\}$ *and some time steps* $k_{i,t} \in [k_0, k_l - 1]$ *and* $k_{i,\bar{t}} \in [k_0, k_l - 1]/\{k_{i,t}\}$ *such that*

$$|\varrho_i^{l,l'}(x(k_{i,t}), u(k_{i,t}))| > (q_i - 1)\varepsilon_i^*,$$
$$|\varrho_i^{l,l'}(x(k_{i,\bar{t}}), u(k_{i,\bar{t}}))| < (q_i - 1)\varepsilon_i^*, \tag{21}$$

*with the total number of the elements in set* $\{k_{i,t}\}$ *(denoted by* $T_i$*) satisfying* $0 \le T_i < \log_{b_i} \frac{\gamma_i-(q_i-1)\varepsilon_i^*}{\gamma_i-\mu_i}$, *where* $\mu_i := \max_{k \in \{k_{i,\bar{t}}\}}\{|\varrho_i^{l,l'}(x(k), u(k))|\}$ *and* $\gamma_i := \max_{k \in \{k_{i,t}\}}\{|\varrho_i^{l,l'}(x(k), u(k))|\}$.

2) for each mismatched fault $\bar{l}$ ($\bar{l} \in \{1, \cdots, N\}/\{l\}$), condition (18) is satisfied;

then, we have $e_i^l(k_l) < \bar{e}_i(k_l)$ and $e_i^{\bar{l}}(k_l) > \bar{e}_i(k_l)$, i.e., the occurring fault $l'$ is isolated at time $k_l$. The upper bound on the isolation time $K_l = k_l - k_0$ is given by

$$\bar{K}_l = \max_{i=1,2,\cdots,n} \{\lceil \log_{b_i} \frac{b_i^{T_i}(\gamma_i - \mu_i) + (q_i - 1)\varepsilon_i^* - \gamma_i}{(1 - b_i)E_i^l + (q_i - 1)\varepsilon_i^* - \mu_i} \rceil\},$$
(22)

with $E_i^l = \max_{k \le k_0}\{e_i^l(k)\}$.

*Proof:*

For the matched fault $l$, under condition (21), we have $|\varrho_i^{l,l'}(x(k_{i,t}), u(k_{i,t}))| \le \gamma_i$ for all $k_{i,t} \in [k_0, k_l - 1]$, and $|\varrho_i^{l,l'}(x(k_{i,\bar{t}}), u(k_{i,\bar{t}}))| \le \mu_i$ for all $k_{i,\bar{t}} \in [k_0, k_l - 1]/\{k_{i,t}\}$. Recalling that for $k > k_0$, $(x, u) = (x^l, u^l)$, and $|\epsilon_i^l| = |\epsilon_{i,1}^l| \le \varepsilon_i^*$ according to (10), and noting that $E_i^l \ge e_i^l(k_0)$, the state of the $l$-th residual system in (14) satisfies

$e_i^l(k_l)$

$= b_i^{k_l - k_0} e_i^l(k_0) + \sum_{h=k_0}^{k_l-1} b_i^{k_l-1-h} |\varrho_i^{l,l'}(x(h), u(h)) - \epsilon_i^l|$

$\le b_i^{k_l - k_0} e_i^l(k_0) + \sum_{h=k_0}^{k_l-1} b_i^{k_l-1-h} (|\epsilon_i^l| + |\varrho_i^{l,l'}(x(h), u(h))|)$

$\le b_i^{k_l - k_0} E_i^l + \sum_{h=k_0}^{k_l-1} b_i^{k_l-1-h} (\varepsilon_i^* + |\varrho_i^{l,l'}(x(h), u(h))|)$

$= b_i^{k_l - k_0} E_i^l + \sum_{h=k_{i,t}} b_i^{k_l-1-h} |\varrho_i^{l,l'}(x(h), u(h))|$

$+ \sum_{h=k_{i,\bar{t}}} b_i^{k_l-1-h} |\varrho_i^{l,l'}(x(h), u(h))| + \frac{(1 - b_i^{k_l - k_0})\varepsilon_i^*}{1 - b_i}$

$\le b_i^{k_l - k_0} E_i^l + \frac{(1 - b_i^{k_l - k_0})\varepsilon_i^*}{1 - b_i} + \sum_{h=k_{i,t}} b_i^{k_l-1-h} \gamma_i$

$+ \sum_{h=k_{i,\bar{t}}} b_i^{k_l-1-h} \mu_i$

$= b_i^{k_l - k_0} E_i^l + \frac{(1 - b_i^{k_l - k_0})\varepsilon_i^*}{1 - b_i} + \sum_{h=k_{i,t}} b_i^{k_l-1-h} (\gamma_i - \mu_i)$

$+ \sum_{h=k_{i,t}} b_i^{k_l-1-h} \mu_i + \sum_{h=k_{i,\bar{t}}} b_i^{k_l-1-h} \mu_i$

$= b_i^{k_l - k_0} E_i^l + \frac{(1 - b_i^{k_l - k_0})\varepsilon_i^*}{1 - b_i} + \sum_{h=k_{i,t}} b_i^{k_l-1-h} (\gamma_i - \mu_i)$

$+ \sum_{h=k_0}^{k_l-1} b_i^{k_l-1-h} \mu_i$

$= b_i^{k_l - k_0} E_i^l + \frac{(1 - b_i^{k_l - k_0})\varepsilon_i^*}{1 - b_i} + \sum_{h=k_{i,t}} b_i^{k_l-1-h} (\gamma_i - \mu_i)$

$+ \frac{1 - b_i^{k_l - k_0}}{1 - b_i} \mu_i,$
(23)

where $\sum_{h=k_{i,t}} b_i^{k_l-1-h} |\varrho_i^{l,l'}(x(h), u(h))|$ is the sum of $b_i^{k_l-1-k_{i,t}} |\varrho_i^{l,l'}(x(k_{i,t}), u(k_{i,t}))|$ for $T_i$ numbers of $k_{i,t}$, and $\sum_{h=k_{i,\bar{t}}} b_i^{k_l-1-h} |\varrho_i^{l,l'}(x(h), u(h))|$ is the sum of $b_i^{k_l-1-k_{i,\bar{t}}} |\varrho_i^{l,l'}(x(k_{i,\bar{t}}), u(k_{i,\bar{t}}))|$ for $k_l - k_0 - T_i$ numbers of $k_{i,\bar{t}}$.

From condition (21), we know that $\gamma_i > (q_i - 1)\varepsilon^* > \mu_i$, and thus $\frac{\gamma_i - (q_i - 1)\varepsilon_i^*}{\gamma_i - \mu_i} > 1$. Therefore existence of $T_i$ can be guaranteed. Based on this, noting that $0 \le b_i < 1$, it can be verified that the term $\sum_{h=k_{i,t}} b_i^{k_l-1-h}(\gamma_i - \mu_i)$ in (23) satisfies

$$\sum_{h=k_{i,t}}^{} b_i^{k_l-1-h}(\gamma_i - \mu_i) \le \sum_{h=k_l-T_i}^{k_l-1} b_i^{k_l-1-h}(\gamma_i - \mu_i)$$
$$= \frac{1 - b_i^{T_i}}{1 - b_i}(\gamma_i - \mu_i),$$
(24)

which results in

$$e_i^l(k_l) \le b_i^{k_l - k_0} E_i^l + \frac{(1 - b_i^{k_l - k_0})\varepsilon_i^*}{1 - b_i}$$
$$+ \frac{1 - b_i^{T_i}}{1 - b_i}(\gamma_i - \mu_i) + \frac{1 - b_i^{k_l - k_0}}{1 - b_i}\mu_i.$$
(25)

From (25), $e_i^l(k_l) \le \frac{(1 - b_i^{k_l - k_0})q_i\varepsilon_i^*}{1 - b_i}$ holds as long as

$$K_l = k_l - k_0 \ge \log_{b_i} \frac{b_i^{T_i}(\gamma_i - \mu_i) + (q_i - 1)\varepsilon_i^* - \gamma_i}{(1 - b_i)E_i^l + (q_i - 1)\varepsilon_i^* - \mu_i},$$
(26)

where $(1 - b_i)E_i^l + (q_i - 1)\varepsilon_i^* - \mu_i > (q_i - 1)\varepsilon_i^* - \mu_i > 0$ and $b_i^{T_i}(\gamma_i - \mu_i) + (q_i - 1)\varepsilon_i^* - \gamma_i > 0$ given the results of $\gamma_i > (q_i - 1)\varepsilon_i^* > \mu_i$ and $T_i < \log_{b_i} \frac{\gamma_i - (q_i - 1)\varepsilon_i^*}{\gamma_i - \mu_i}$ under the condition (21). Furthermore, we have

$$\log_{b_i} \frac{b_i^{T_i}(\gamma_i - \mu_i) + (q_i - 1)\varepsilon_i^* - \gamma_i}{(1 - b_i)E_i^l + (q_i - 1)\varepsilon_i^* - \mu_i}$$
$$> \log_{b_i} \frac{b_i^{T_i}(\gamma_i - \mu_i) + b_i^{T_i}((q_i - 1)\varepsilon_i^* - \gamma_i)}{(1 - b_i)E_i^l + (q_i - 1)\varepsilon_i^* - \mu_i}$$
$$> \log_{b_i} \frac{b_i^{T_i}((1 - b_i)E_i^l + (q_i - 1)\varepsilon_i^* - \mu_i)}{(1 - b_i)E_i^l + (q_i - 1)\varepsilon_i^* - \mu_i} = T_i,$$

which implies that (26) ensures $K_l = k_l - k_0 > T_i$. Consequently, $e_i^l(k_l) \le \frac{(1-b_i^{k_l-k_0})q_i\varepsilon_i^*}{1-b_i} < \frac{(1-b_i^{k_l})q_i\varepsilon_i^*}{1-b_i} = \bar{e}_i(k_l)$ is guaranteed under condition (21). Furthermore, we have $e_i^{\bar{l}}(k_l) > \bar{e}_i(k_l)$ under the condition (18), as verified in Theorem 1. Thus, it can be deduced that the occurring fault $l'$ will be identified similar to the trained fault $l$ in a finite time $k_l$.

Finally, since $K_l \in \mathbb{Z}_+$ and according to (26), the upper bound of the isolation time $\bar{K}_l$ can be obtained as:

$$\bar{K}_l = \max_{i=1,2,\cdots,n} \{\lceil \log_{b_i} \frac{b_i^{T_i}(\gamma_i - \mu_i) + (q_i - 1)\varepsilon_i^* - \gamma_i}{(1 - b_i)E_i^l + (q_i - 1)\varepsilon_i^* - \mu_i} \rceil\},$$
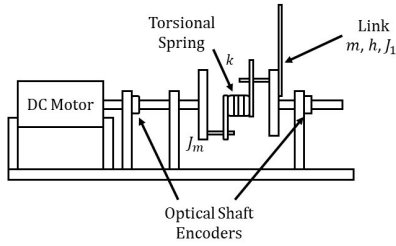(27)

which ends the proof. □

**FIGURE 3.** Schematic of an elastic robot.

*Remark 9: In the above theorem, condition (21) implies that during $[k_0, k_l - 1]$, an isolatable fault $l'$ is required to sufficiently match the trained fault $l$ for most of time (denoted by $k_{i,\bar{t}}$), i.e., the associated fault difference $|\varrho_i^{l,l'}(x, u)|$ needs to be relatively small. On the other hand, it is also allowable that there exist some time instants $k_{i,t}$ (which could be non-consecutive) when the fault difference is not sufficiently small (while its upper bound still exists as denoted by $\gamma_i$). Theorem 2 further implies that if the parameters $\mu_i$, $\gamma_i$ and $T_i$ have relatively small values, i.e., the occurring fault $l'$ perfectly matches the fault $l$, the isolation time $\bar{K}_l$ in (22) then becomes very small, indicating that fault isolation could be achieved in a rapid manner.*

*Remark 10: In Theorems 1 and 2, the effect of the parameter $q_i$ on the isolation performance has been established. Specifically, from (27), it is seen that increasing $q_i$ would decrease the value of $\bar{K}_l$, enabling the isolation to be achieved in a rapid manner. On the other hand, increasing $q_i$ will increase the bounds given in (17)–(18), which could result in failure of achieving accurate isolation for those similar faults with relatively small mutual differences. Thus, when designing the value for $q_i$, the tradeoff between isolation time and isolation accuracy should be taken into account. An example will be given in next section for illustration.*

## IV. SIMULATION RESULTS
In this section, to verify the effectiveness and advantages of our scheme, we consider a sFI problem for a single-link flexible joint robot [32], [33] as shown in Fig. 3. Its motion equations are described as

$$J_1 \ddot{q}_1 + k(q_1 - q_2) + mgh \sin q_1 = 0$$
$$J_m \ddot{q}_2 + F_m \dot{q}_2 - k(q_1 - q_2) = k_\tau u, \qquad (28)$$

where $q_1$ and $q_2$ denote the angular positions of the controlled link and the motor, respectively; $u$ is the control torque generated by the motor; $J_m = 3.7 \times 10^{-3} kg \cdot m^2$ is the inertia of the motor; $J_1 = 9.3 \times 10^{-3} kg \cdot m^2$ is the inertia of the link; the torsional spring constant is $k = 1.8 \times 10^{-1} Nm/rad$; link mass is $m = 2.1 \times 10^{-1} kg$; link length is $h = 0.5m$; amplifier gain is $k_\tau = 8 \times 10^{-2} Nm/V$; viscous friction coefficient is $F_m = 4.6 \times 10^{-2} Nm/V$; and the gravity constant is $g = 9.8 \ m/s^2$.

Choosing $x_1 = q_1, x_2 = \dot{q}_1, x_3 = q_2, x_4 = \dot{q}_2$, and through Euler approximation, the system (28) can be discretized as:

$$
\begin{aligned}
x_1(k+1) &= x_1(k) + T_s x_2(k) \\
x_2(k+1) &= x_2(k) + \frac{T_s}{J_1}(-mgh \sin(x_1(k)) \\
&\quad - k(x_1(k) - x_3(k))) \\
x_3(k+1) &= x_3(k) + T_s x_4(k) \\
x_4(k+1) &= x_4(k) + \frac{T_s}{J_m}(k(x_1(k) - x_3(k)) \\
&\quad - F_m x_4(k) + k_\tau u(k)), \qquad (29)
\end{aligned}
$$

with $T_s = 0.01s$ being the sampling period. For simulation purpose, the system input is given by $u(k) = 2 \sin(2T_s k)$, and the system dynamics in (29) is assumed completely unknown, i.e., $v_1(x(k), u(k)) := x_1(k) + T_s x_2(k)$, $v_2(x(k), u(k)) := x_2(k) + \frac{T_s}{J_1}(-mgh \sin(x_1(k)) - k(x_1(k) - x_3(k)))$, $v_3(x(k), u(k)) := x_3(k) + T_s x_4(k)$ and $v_4(x(k), u(k)) := x_4(k) + \frac{T_s}{J_m}(k(x_1(k) - x_3(k)) - F_m x_4(k) + k_\tau u(k))$ are all unknown functions. The following four types of faults are considered for simulation purpose:

- Fault 1: A multiplicative actuator fault that is modeled by letting $u'(k) = (1 + \theta_1)u(k)$, where $u$ is the normal control input when no fault occurs and $\theta_1$ is the parameter characterizing the magnitude of the fault. Note that the case of $\theta_1 = 0$ represents the normal operation condition (no fault). The associated fault function can be described by $\phi_4^1(x(k), u(k)) := \frac{T_s k_\tau \theta_1}{J_m} u(k)$. We set $\theta_1 = 0.2$ for simulation purpose.
- Fault 2: An oscillation actuator fault by letting $u'(k) = u(k) + \theta_1(\theta_2 \sin(20 T_s(k - 1)) + 1)u(k)$, where $\theta_2$ is the parameter describing the level of oscillation. We denote fault 2 as $\phi_4^2(x(k), u(k)) := \frac{T_s k_\tau \theta_1}{J_m}(\theta_2 \sin(20 T_s(k-1)) + 1)u(k)$ with $\theta_2 = 0.4$.
- Fault 3: A fault leading to extra abnormal friction in the motor by letting $F'_m = F_m + F_{mf_1}$, where $F_m$ is the normal viscous friction constant, $F_{mf_1}$ is the extra abnormal viscous friction constant. In this case, the fault function is $\phi_4^3(x(k), u(k)) := -\frac{T_s F_{mf_1}}{J_m} x_4(k)$, we set $F_{mf_1} = -0.02$ for simulation.
- Fault 4: An oscillation fault occurring in the motor by letting $F'_m = F_m + F_{mf_1}(F_{mf_2} \cos(20 T_s(k - 1) + 1)$, with $F_{mf_2}$ characterizing the strength of the oscillation. The fault function is denoted by $\phi_4^4(x(k), u(k)) := -\frac{T_s F_{mf_1}}{J_m}(F_{mf_2} \cos(20 T_s(k-1)) + 1)x_4(k)$. We set $F_{mf_2} = 0.3$.

For the monitored system (29), the system states $(x_1, x_2, x_3, x_4)$ under all possible modes (including normal mode and four faulty modes) are plotted in Fig. 4, indicating that all faulty system trajectories are close to each other. The fault mismatch functions, i.e., $\varrho_4^{1,2}(x, u) = \phi_4^1(x, u) - \phi_4^2(x, u)$, $\varrho_4^{2,3}(x, u) = \phi_4^2(x, u) - \phi_4^3(x, u)$, $\varrho_4^{3,4}(x, u) = \phi_4^3(x, u) - \phi_4^4(x, u)$, $\varrho_4^{4,1}(x, u) = \phi_4^4(x, u) - \phi_4^1(x, u)$, are compared with the system uncertainty $v_4(x, u)$ in Fig. 5. It is seen that the magnitude of each fault mismatch function is noticeably smaller compared to that of the system uncertainty.
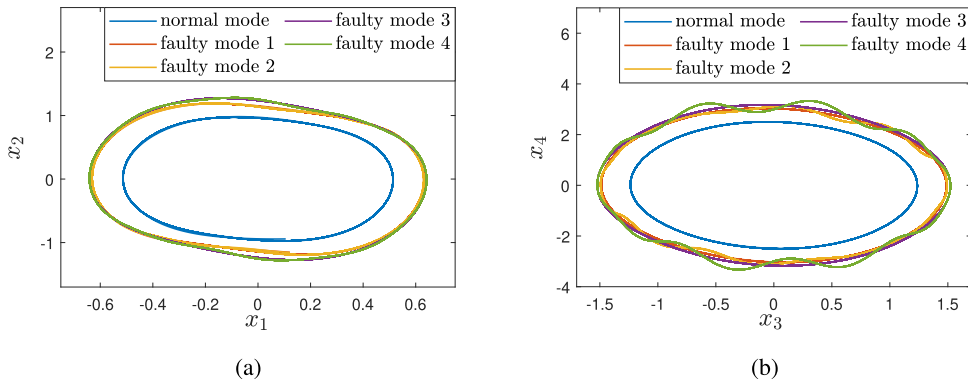
**FIGURE 4.** System state trajectories under normal mode and all possible faulty modes. (a) $(x_1, x_2)$. (b) $(x_3, x_4)$.
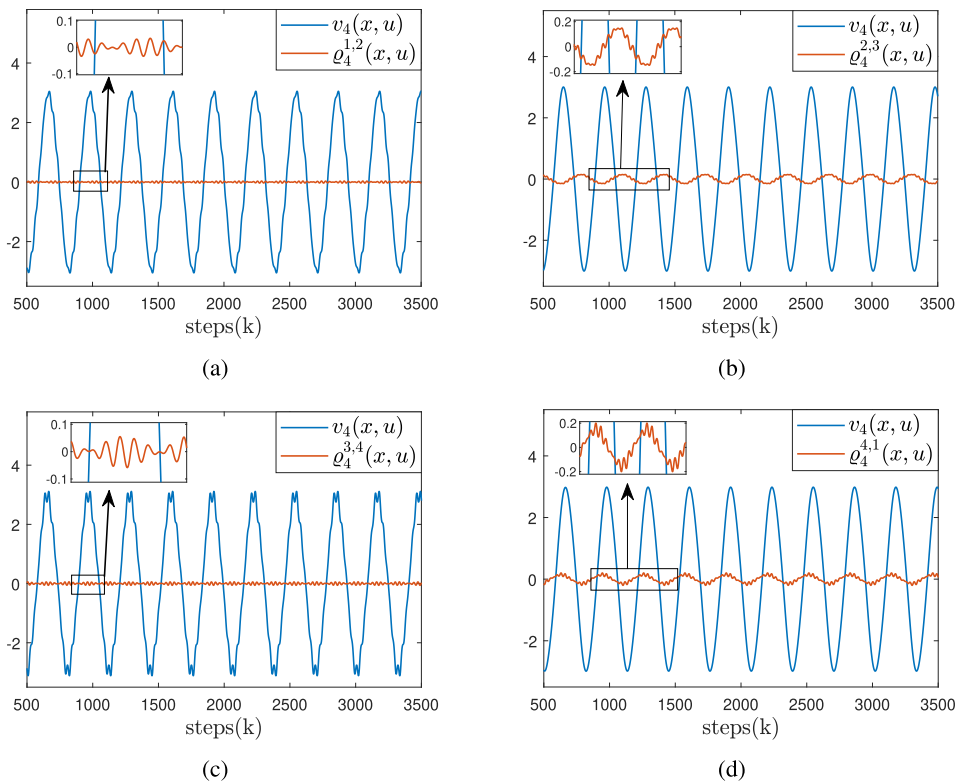


**FIGURE 5.** Comparison between system uncertainty $v_4(x, u)$ and the fault mismatch functions $\varrho_4^{s,l}(x, u) = \phi_4^s(x, u) - \phi_4^l(x, u)$ $(s, l \in \{1, 2, 3, 4\})$. (a) $v_4(x, u)$ and $\varrho_4^{1,2}(x, u)$. (b) $v_4(x, u)$ and $\varrho_4^{2,3}(x, u)$. (c) $v_4(x, u)$ and $\varrho_4^{3,4}(x, u)$. (d) $v_4(x, u)$ and $\varrho_4^{4,1}(x, u)$.

This demonstrates that the considered four types of faults are similar faults.

Given the above system setup, to examine the effectiveness of our sFI scheme, we first achieve the accurate identification for the system dynamics under all faulty modes. For each $s$-th faulty mode ($s = 1, 2, 3, 4$), the proposed identifier consisting of (6) and (7) is implemented to learn the unknown dynamics $x_i(k + 1) = v_i(x(k), u(k)) + \phi_i^s(x(k), u(k))$ ($i = 1, 2, 3, 4$) in (29). The desired learning accuracy level is pre-designed as $\varepsilon_1^* = 0.0082$, $\varepsilon_2^* = 0.0293$, $\varepsilon_3^* = 0.0215$, $\varepsilon_4^* = 0.0571$, which can be achieved by constructing

a sufficiently large number of neurons. Through extensive trial and error, we construct the RBF networks $\hat{W}_i^{sT} S(x, u)$ in a regular lattice, with nodes $N = 7 \times 11 \times 13 \times 23 \times 14$, the centers evenly spaced on $[-0.9, 0.9] \times [-1.5, 1.5] \times [-1.8, 1.8] \times [-3.3, 3.3] \times [-2, 2]$ and the widths $\eta_t = 0.3$. The weights of the RBF networks $\hat{W}_i^s$ are updated according to (7). The design parameters of (6) and (7) are $a_i = 0.2$ and $c_i = 0.1$. The initial conditions are set as $\hat{W}_i^s(0) = 0$, $u(0) = 0$, $x(0) = [0, 0, 0, 0]^T$ and $\hat{x}(0) = [0, 0, 0, 0]^T$. All the above system setups are kept the same for all $i = 1, 2, 3, 4$ and $s = 1, 2, 3, 4$. In the following, we only show the identification
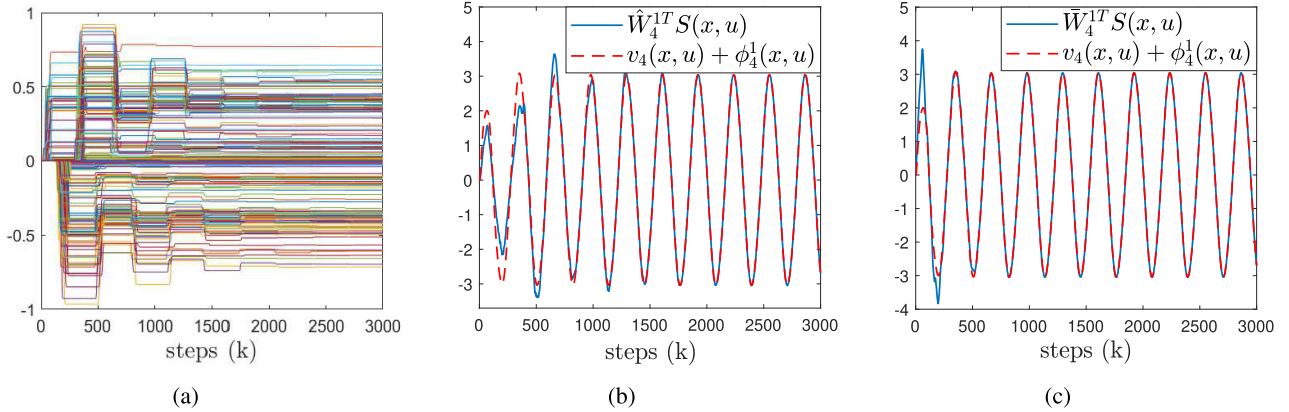
**FIGURE 6.** Identification of the unknown function $v_4(x, u) + \phi_4^1(x, u)$ in faulty mode 1. (a) Weight convergence of $\hat{W}_4^1$. (b) Function approximation: $\hat{W}_4^{1T} S(x, u)$ and $v_4(x, u) + \phi_4^1(x, u)$. (c) Function approximation: $\bar{W}_4^{1T} S(x, u)$ and $v_4(x, u) + \phi_4^1(x, u)$.
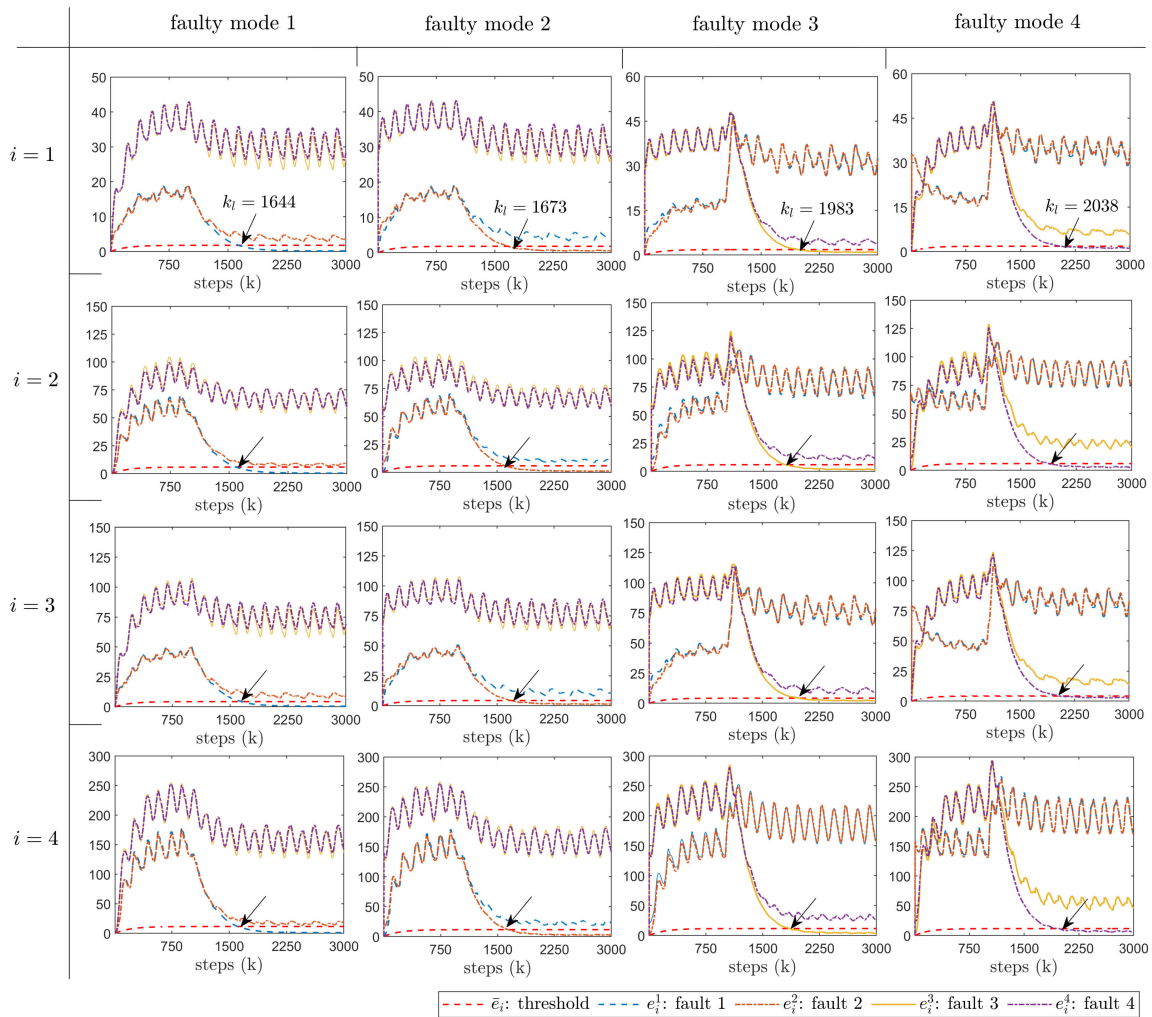


**FIGURE 7.** Fault isolation results under different faulty modes ($s = 1, 2, 3, 4$) by monitoring the time profiles of residual signals $e_i^s$ in (30) for all $i = 1, 2, 3, 4$ and the associated threshold signals $\bar{e}_i$ in (16) for all $i = 1, 2, 3, 4$.

results of the unknown function $v_4(x, u) + \phi_4^1(x, u)$ in faulty mode 1. Simulation results for learning the faulty modes 2, 3 and 4 are similar and omitted here due to the limitation of space. Specifically, Fig. 6a shows the convergence

performance of weights $\hat{W}_4^1$. The accurate approximation for the function $v_4(x, u) + \phi_4^1(x, u)$ is achieved by the constructed RBF NN $\hat{W}_4^{1T} S(x, u)$ as shown in Fig. 6b. Thanks to the convergence of $\hat{W}_4^1$, the constant model $\bar{W}_4^{1T} S(x, u)$ is
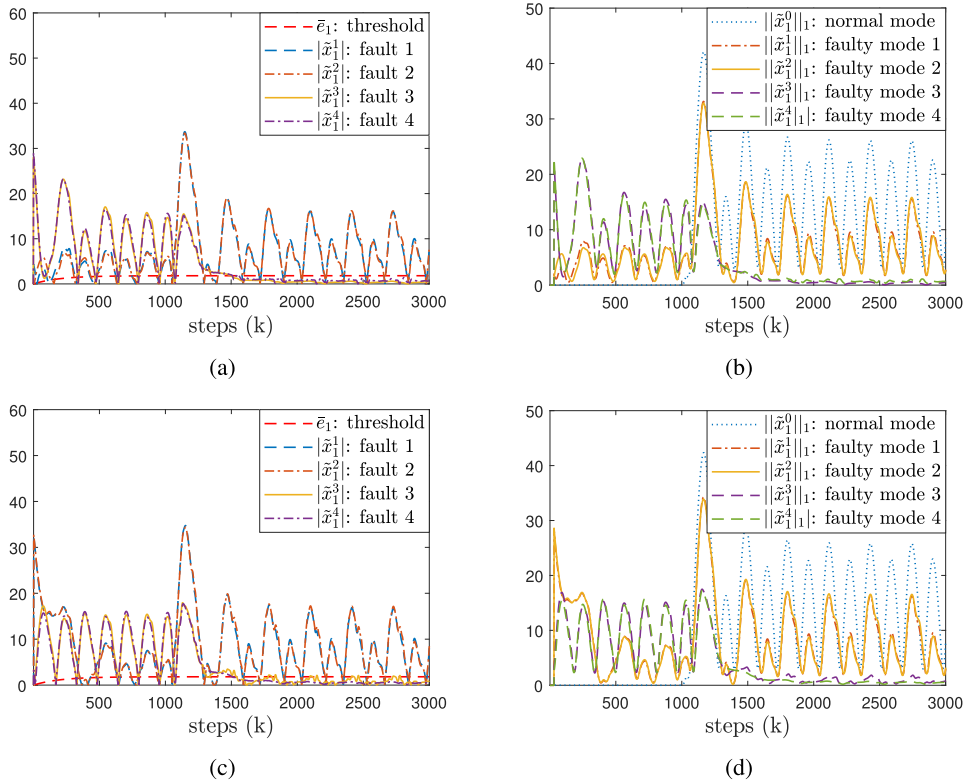
**FIGURE 8.** Fault isolation using the methods [15], [16], [18], [19] with utilization of direct measurement of system dynamics differences. (a) Isolation of fault 3′ with methods [15], [16]. (b) Isolation of fault 3′ with methods [18], [19]. (c) Isolation of fault 4′ with methods [15], [16]. (d) Isolation of fault 4′ with methods [18], [19].

obtained by $\bar{W}_4^1 = \frac{1}{100} \sum_{k=2901}^{3000} \hat{W}_4^1(k)$, which can be used to represent the unknown dynamics $v_4(x, u) + \phi_4^1(x, u)$, as shown in Fig. 6c.

Using the constant models $\bar{W}_i^{sT} S(x, u)$ $(i, s = 1, 2, 3, 4)$, from (12), we construct the following residual systems:

$$e_i^s(k) = b_i e_i^s(k-1) + |\bar{W}_i^{sT} S(x(k-1), u(k-1)) - x_i(k)|, \tag{30}$$

with $b_i = 0.995$. The adaptive threshold $\bar{e}_i$ is implemented according to (16) by setting $q_1 = 1.1$, $q_2 = 1.01$, $q_3 = 1.01$ and $q_4 = 1.01$. For simulation purpose, the tested faults are designed as: fault 1′ (similar to the trained fault 1) with $\phi_4^{1'}(x, u) := \frac{T_s k_\tau \theta_1'}{J_m} u$ $(\theta_1' = 0.201)$, fault 2′ (similar to the trained fault 2) with $\phi_4^{2'}(x, u) = \frac{T_s k_\tau \theta_1'}{J_m}(\theta_2' \sin(20\, T_s(k-1)) + 1)u$ $(\theta_2' = 0.42)$, fault 3′ (similar to the trained fault 3) with $\phi_4^{3'}(x, u) := -\frac{T_s F_{mf_1}'}{J_m} x_4$ $(F_{mf_1}' = -0.0205)$, and fault 4′ (similar to the trained fault 4) with $\phi_4^{4'}(x, u) := -\frac{T_s F_{mf_1}'}{J_m}(F_{mf_2}' \cos(20\, T_s(k-1)) + 1)x_4$ $(F_{mf_2}' = 0.302)$. The fault occurrence time for these tested faults is set as $k_0 = 1000$. The sFI simulation results for these tested faults are plotted in Fig. 7. Considering the case of faulty mode 1, i.e., the fault 1′ with $\theta_1' = 0.201$ occurs in the monitored system (29), it can be seen from Fig. 7 that the matched residual $e_i^1$ $(i = 1, 2, 3, 4)$ becomes smaller than the corresponding

threshold $\bar{e}_i$ at time $k_l = 1644$, whereas other mismatched residuals $e_i^s$ $(s = 2, 3, 4, i = 1, 2, 3, 4)$ remain no smaller than the thresholds. Thus fault 1′ is isolated at $k_l = 1644$, with an isolation time $K_l = k_l - k_0 = 644$. Similarly, it is also observed in Fig. 7 that the fault 2′ is isolated at $k_l = 1673$, fault 3′ at $k_l = 1983$ and fault 4′ at $k_l = 2038$.

To verify the superiority of our sFI scheme over existing approaches, we conduct simulation comparisons with existing approaches of the adaptive threshold mechanism-based FI scheme [15], [16] and the SRP-based FI schemes [18], [19]. All these approaches derive the residuals through a mechanism of direct measurements of system dynamics differences. Specifically, from [15], [16], the estimators for each trained faulty mode $s$ $(s = 1, 2, 3, 4)$ are constructed as:

$$\bar{x}_i^s(k) = b_i(\bar{x}_i^s(k-1) - x_i(k-1)) + \bar{W}_i^{sT} S(x(k-1), u(k-1)), \quad i = 1, 2, 3, 4, \tag{31}$$

where $\bar{W}_i^{sT} S(x, u)$ is the constant network obtained from the learning phase as discussed above, $b_i$ is set as 0.995. By comparing the dynamics of FI systems (31) with that of monitored system (29), the residuals used for FI are obtained as $|\tilde{x}_i^s| = |\bar{x}_i^s - x_i|$. In another class of FI schemes [18], [19], the estimators are constructed for not only the faulty modes $s = 1, 2, 3, 4$ but also the normal mode $s = 0$, which can be implemented following (31). The residual
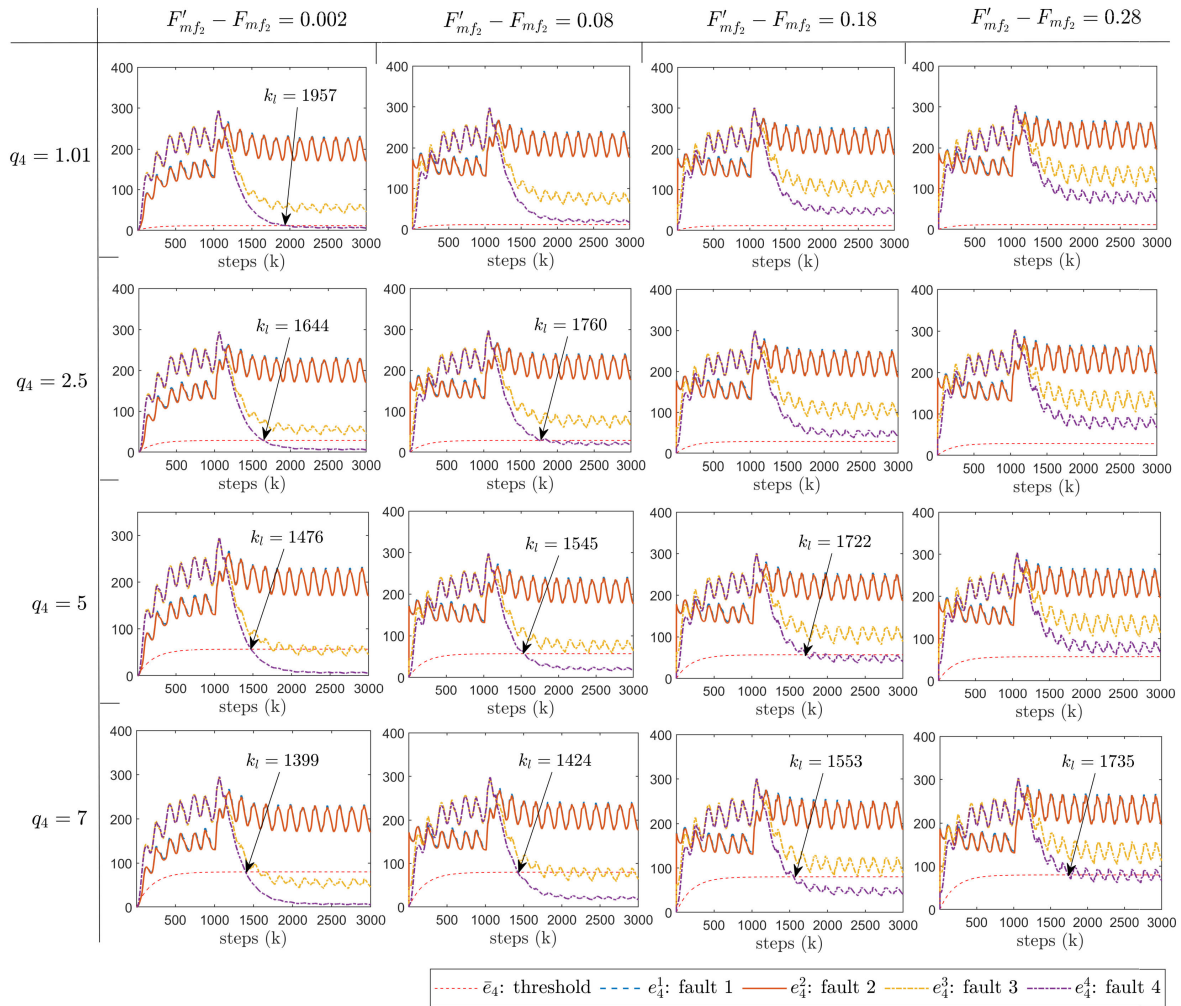
**FIGURE 9.** Fault isolation results with different tested faults 4′ and different parameters $q_4$. $F'_{mf_2} - F_{mf_2}$ denotes the difference between the tested fault 4′ and the matched fault 4.

signals are generated in real time using an average $L_1$ norm, i.e., $||\tilde{x}_i^s||_1 = ||\tilde{x}_i^s - x_i||_1 = \frac{1}{K}\sum_{h=k-K}^{k-1}|\tilde{x}_i^s(h) - x_i(h)|$, where the coefficient $K$ is selected as $K = 30$. With the methods [15], [16] and [18], [19], for fault 3′ (as defined above with $F'_{mf_1} = -0.0205$), the FI is achieved in Figs. 8a and 8b, respectively. It is seen that after fault occurrence, the mismatched residuals $|\tilde{x}_1^4|$ (of [15], [16]) and $||\tilde{x}_1^4||_1$ (of [18], [19]) will approach zero and finally be mixed up with the matched residuals, i.e., $|\tilde{x}_1^3|$ and $||\tilde{x}_1^3||_1$. This implies that the fault difference between the faults 3′ and 4 cannot be accurately captured by the residuals $|\tilde{x}_1^4|$, $||\tilde{x}_1^4||_1$. This is because the fault difference between the 3-rd fault type and 4-th fault type, i.e., $\varrho_4^{3,4}(x, u)$ shown in Fig. 5c, has frequently-changing signs. With the residual systems of [15], [16] and [18], [19], the accumulated effect of $\varrho_4^{3,4}(x, u)$ will be offset and cannot manifest itself in the residuals $|\tilde{x}_1^4|$, $||\tilde{x}_1^4||_1$. As a result, under the approaches [15], [16] and [18], [19], the fault 3′ could be identified similar to fault 3 as well as fault 4. Similar observation can also be found when fault 4′ (as defined above

with $F'_{mf_2} = 0.302$) occurs, as shown in Fig. 8c. These comparison results justify the superiority of our approach over the FI schemes of [15], [16], [18], [19].

In the following, we further investigate the relationship between the parameter $q_i$ of the adaptive threshold (16) and the isolation performance (in terms of isolation accuracy and isolation time). We consider a bank of tested faults that are similar to the fault 4 (with $\phi_4^4(x, u) = -\frac{T_s F_{mf_2}}{J_m}(F_{mf_2}\cos(20\ T_s(k - 1)) + 1)x_4$, $F_{mf_1} \equiv -0.02$, $F_{mf_2} \equiv 0.3$), i.e., $\phi_4^{4'}(x, u) = -\frac{T_s F'_{mf_1}}{J_m}(F'_{mf_2}\cos(20\ T_s(k - 1)) + 1)x_4$ with $F'_{mf_1} \equiv -0.02$. The parameter $F'_{mf_2}$ are selected as $F'_{mf_2} = 0.302, 0.38, 0.48, 0.58$ (or $F'_{mf_2} - F_{mf_2} = 0.002, 0.08, 0.18, 0.28$), representing four tested faults with various fault differences. The fault occurrence time is assumed at $k_0 = 1000$. The simulation study is carried out by setting the parameter $q_i$ in different values, i.e., $q_i = 1.01, 2.5, 5, 7$. The other system setups including the design of residual systems (30) and the parameters $b_i, \varepsilon_i^*$
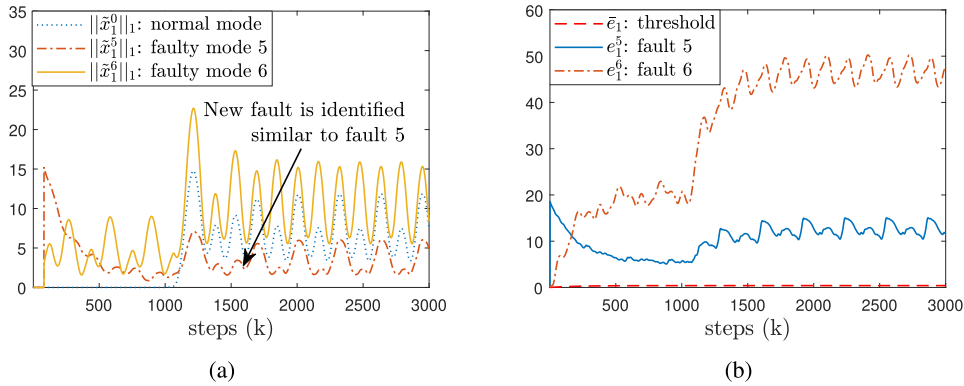
**FIGURE 10.** Isolation of a new fault 7 using different methods. (a) SRP-based method in [18], [19]. (b) Our proposed method based on adaptive threshold mechanism.

$(i = 1, 2, 3, 4)$ in (16) are kept the same as above. Due to the limitation of space, only the simulation results of $i = 4$ is presented (as shown in Fig. 9), the results of $i = 1, 2, 3$ are similar thus omitted here. We first consider the case of $q_4 = 1.01$. In Fig. 9, only the tested fault with $F'_{mf_2} - F_{mf_2} = 0.002$ is isolatable (identified similar to trained fault 4), whereas those tested faults with $F'_{mf_2} - F_{mf_2} \geq 0.08$ will be identified as new faults. When increasing $q_4$, as shown in Fig. 9, those tested faults with larger fault differences, i.e., the magnitude of $F'_{mf_2} - F_{mf_2}$ becomes larger, can be isolated, and the FI time $k_I$ is decreased. When the value $q_4$ becomes too large, e.g., $q_4 = 5$ and $q_4 = 7$, some tested faults (e.g., with $F'_{mf_2} - F_{mf_2} = 0.002$, and $F'_{mf_2} - F_{mf_2} = 0.08$) would be identified to match both the fault 4 and fault 3. In such cases, the 3-rd fault type and 4-th fault type will be considered as the same type of fault. These simulation results demonstrate the effect of the parameter $q_i$ on the isolation performance, which illustrate the discussions in Remark 10.

Finally, we further verify that our scheme has better isolation reliability compared to the existing SRP-based FI methods of [18], [19]. To this end, two types of faults will be considered for training: fault 5: $\phi_4^5(x, u) := \frac{T_s k_\tau \theta_5}{J_m} u$ with $\theta_5 = 0.1$, and fault 6: $\phi_4^6(x, u) := -\frac{T_s \theta_6}{J_m} x_4$ with $\theta_6 = 0.01$. Specific implementation of our method and the method in [18], [19] follows the procedure as detailed in (30) and (31), respectively. The associated parameters for our method are set as $\varepsilon_1^* = 0.0018$, $\varepsilon_2^* = 0.0056$, $\varepsilon_3^* = 0.0045$, $\varepsilon_4^* = 0.0117$ and $b_i = 0.995$, $q_i = 1.01$ $(i = 1, 2, 3, 4)$, and the parameters for method [18], [19] are set as $b_i = 0.995$ and $K = 80$. For comparison, we assume a new fault (i.e., fault 7: $\phi_4^7(x, u) := -\frac{T_s \theta_7}{J_m} (\theta_8 \sin(20\, T_s (k-1)) + 1) x_3$ with $\theta_7 = -0.02$, $\theta_8 = 0.3$) occurs in the system (29). The isolation results using the method [18], [19] and our method are illustrated in Figs. 10a and 10b, respectively. It can be seen that under the SRP-based methods [18], [19] (see, Fig. 10a), fault 7 is identified similar to fault 5, which is an isolation misjudgment. Nevertheless, under our method (see, Fig. 10b), the fault 7 is identified as a new fault that does not match any trained faults. These comparisons justify our discussions in Remark 5.

## V. CONCLUSIONS

In this paper, we have proposed a novel adaptive threshold based sFI scheme for discrete-time nonlinear uncertain systems. The scheme consists of the fault dynamics identification phase, and the fault isolation phase. In identification phase, a DL-based adaptive dynamics learning approach has been employed to locally-accurately identify uncertain system dynamics. The learned knowledge can be obtained and stored in a bank of constant RBF NNs models. In fault isolation phase, with the obtained knowledge, a bank of sFI estimators (residual systems) have been developed through a novel mechanism of absolute measurement of fault dynamics differences. Their generated residuals are able to effectively capture the small fault differences and distinguish them from the system uncertainty. An adaptive threshold was then proposed based on the designed sFI estimators, such that the sFI decision making can be achieved by comparing the residuals with the designed threshold in real time. Rigorous analysis on isolatability condition and isolation time has been conducted to characterize the isolation capabilities of our scheme. The effectiveness of all the proposed results has been verified by extensive simulation studies on a practical application of a single-link robotic arm.

For the future work, we plan to extend the current results to (i) those cases when full system information is not available (e.g., only partial system states or measurement outputs are available) for fault isolation designs, promising approaches including the state observer techniques from [35], [36] will be explored; and (ii) applications of the proposed approach to fault isolation of real engineering systems, such as robotic manipulators or unmanned vehicles, etc.

## REFERENCES

[1] I. Hwang, S. Kim, Y. Kim, and C. E. Seah, "A survey of fault detection, isolation, and reconfiguration methods," *IEEE Trans. Control Syst. Technol.*, vol. 18, no. 3, pp. 636–653, May 2010.

[2] H. Hamdi, M. Rodrigues, C. Mechmeche, D. Theilliol, and N. B. Braiek, "Fault detection and isolation in linear parameter-varying descriptor systems via proportional integral observer," *Int. J. Adapt. Control Signal Process.*, vol. 26, no. 3, pp. 224–240, Mar. 2012.

[3] R. J. Patton and J. Chen, "Observer-based fault detection and isolation: Robustness and applications," *Control Eng. Pract.*, vol. 5, no. 5, pp. 671–682, May 1997.

[4] A. Mirzaee and K. Salahshoor, "Fault diagnosis and accommodation of nonlinear systems based on multiple-model adaptive unscented Kalman filter and switched MPC and H-infinity loop-shaping controller," *J. Process Control*, vol. 22, no. 3, pp. 626–634, Mar. 2012.

[5] M. Blanke, M. Kinnaert, J. Lunze, M. Staroswiecki, and J. Schroder, *Diagnosis and Fault-Tolerant Control*. Berlin, Germany: Springer, 2006.

[6] A. K. S. Jardine, D. Lin, and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance," *Mech. Syst. Signal Process.*, vol. 20, no. 7, pp. 1483–1510, Oct. 2006.

[7] G. J. Vachtsevanos, F. Lewis, A. Hess, and B. Wu, *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*, vol. 456. Hoboken, NJ, USA: Wiley, 2006.

[8] R. Isermann, R. Schwarz, and S. Stolzl, "Fault-tolerant drive-by-wire systems," *IEEE Control Syst. Mag.*, vol. 22, no. 5, pp. 64–81, Oct. 2002.

[9] X. Zhang, T. Parisini, and M. M. Polycarpou, "Sensor bias fault isolation in a class of nonlinear systems," *IEEE Trans. Autom. Control*, vol. 50, no. 3, pp. 370–376, Mar. 2005.

[10] Q. Zhang, X. An, J. Gu, B. Zhao, D. Xu, and S. Xi, "Application of FBOLES—A prototype expert system for fault diagnosis in nuclear power plants," *Rel. Eng. Syst. Saf.*, vol. 44, no. 3, pp. 225–235, Jan. 1994.

[11] I. A. Gowaid, G. P. Adam, A. M. Massoud, S. Ahmed, D. Holliday, and B. W. Williams, "Quasi two-level operation of modular multilevel converter for use in a high-power DC transformer with DC fault isolation capability," *IEEE Trans. Power Electron.*, vol. 30, no. 1, pp. 108–123, Jan. 2015.

[12] Z. N. S. Vanini, K. Khorasani, and N. Meskin, "Fault detection and isolation of a dual spool gas turbine engine using dynamic neural networks and multiple model approach," *Inf. Sci.*, vol. 259, pp. 234–251, Feb. 2014.

[13] H. C. Cho, J. Knowles, M. S. Fadali, and K. S. Lee, "Fault detection and isolation of induction motors using recurrent neural networks and dynamic Bayesian modeling," *IEEE Trans. Control Syst. Technol.*, vol. 18, no. 2, pp. 430–437, Mar. 2010.

[14] R. M. G. Ferrari, T. Parisini, and M. M. Polycarpou, "A robust fault detection and isolation scheme for a class of uncertain input-output discrete-time nonlinear systems," in *Proc. Amer. Control Conf.*, Jun. 2008, pp. 2804–2809.

[15] R. M. G. Ferrari, T. Parisini, and M. M. Polycarpou, "Distributed fault detection and isolation of large-scale discrete-time nonlinear systems: An adaptive approximation approach," *IEEE Trans. Autom. Control*, vol. 57, no. 2, pp. 275–290, Feb. 2012.

[16] X. Zhang, M. M. Polycarpou, and T. Parisini, "A robust detection and isolation scheme for abrupt and incipient faults in nonlinear systems," *IEEE Trans. Autom. Control*, vol. 47, no. 4, pp. 576–593, Apr. 2002.

[17] X. Zhang, M. M. Polycarpou, and T. Parisini, "Adaptive fault diagnosis and fault-tolerant control of MIMO nonlinear uncertain systems," *Int. J. Control*, vol. 83, no. 5, pp. 1054–1080, May 2010.

[18] T. Chen and C. Wang, "Rapid isolation of small oscillation faults via deterministic learning," *Int. J. Adapt. Control Signal Process.*, vol. 28, nos. 3–5, pp. 366–385, Mar. 2014.

[19] T. Chen, C. Wang, and D. J. Hill, "Rapid oscillation fault detection and isolation for distributed systems via deterministic learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 6, pp. 1187–1199, Jun. 2014.

[20] J. Marzat, H. Piet-Lahanier, F. Damongeot, and E. Walter, "Model-based fault diagnosis for aerospace systems: A survey," *Proc. Inst. Mech. Eng., G, J. Aerosp. Eng.*, vol. 226, no. 10, pp. 1329–1360, Oct. 2012.

[21] E. L. Russell, L. H. Chiang, and R. D. Braatz, *Data-Driven Methods for Fault Detection and Diagnosis in Chemical Process*. London, U.K.: Springer, 2012.

[22] C. Carreras and I. D. Walker, "Interval methods for fault-tree analysis in robotics," *IEEE Trans. Rel.*, vol. 50, no. 1, pp. 3–11, Mar. 2001.

[23] D. A. Bristow and A.-G. Alleyne, "A high precision motion control system with application to microscale robotic deposition," *IEEE Trans. Control Syst. Technol.*, vol. 14, no. 6, pp. 1008–1020, Nov. 2006.

[24] J. Bacher, C. Joseph, and R. Clavel, "Flexures for high precision robotics," *Ind. Robot, Int. J.*, vol. 29, no. 4, pp. 349–353, Aug. 2002.

[25] J. Zhang, C. Yuan, P. Stegagno, H. He, and C. Wang, "Small fault detection of discrete-time nonlinear uncertain systems," *IEEE Trans. Cybern.*, early access, Oct. 22, 2019, doi: 10.1109/TCYB.2019.2945629.

[26] J. Zhang, C. Yuan, and P. Stegagno, "Similar fault isolation of discrete-time nonlinear uncertain systems using smallest residual principle," in *Proc. Amer. Control Conf.*, Denver, CO, USA, Jul. 2020.

[27] J. Park and I. W. Sandberg, "Universal approximation using radial-basis-function networks," *Neural Comput.*, vol. 3, no. 2, pp. 246–257, Jun. 1991.

[28] C. Wang and D. J. Hill, *Deterministic Learning Theory for Identification, Recognition, and Control*. Boca Raton, FL, USA: CRC Press, 2009.

[29] M. Poewell, *The Theory of Radial Basis Function Approximation*. Oxford, U.K.: Clarendon, 1992.

[30] C. Wang and D. J. Hill, "Learning from neural control," *IEEE Trans. Neural Netw.*, vol. 17, no. 1, pp. 130–146, Jan. 2006.

[31] J. Huang, C. Wen, W. Wang, and Y.-D. Song, "Adaptive finite-time consensus control of a group of uncertain nonlinear mechanical systems," *Automatica*, vol. 51, pp. 292–301, Jan. 2015.

[32] M. W. Spong and M. Vidyasagar, *Robot Dynamics and Control*. Hoboken, NJ, USA: Wiley, 1987.

[33] X. Zhang, M. M. Polycarpou, and T. Parisini, "Fault diagnosis of a class of nonlinear uncertain systems with lipschitz nonlinearities using adaptive estimation," *Automatica*, vol. 46, no. 2, pp. 290–299, Feb. 2010.

[34] J. Yao and W. Deng, "Active disturbance rejection adaptive control of uncertain nonlinear systems: Theory and application," *Nonlinear Dyn.*, vol. 89, no. 3, pp. 1611–1624, Aug. 2017.

[35] S. Tong, X. Min, and Y. Li, "Observer-based adaptive fuzzy tracking control for strict-feedback nonlinear systems with unknown control gain functions," *IEEE Trans. Cybern.*, early access, Mar. 18, 2020, doi: 10.1109/TCYB.2020.2977175.

[36] Y.-X. Li and G.-H. Yang, "Observer-based fuzzy adaptive event-triggered control codesign for a class of uncertain nonlinear systems," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 3, pp. 1589–1599, Jun. 2018.

**JINGTING ZHANG** received the B.E. degree from the Department of Automation Science and Engineering, South China University of Technology, China, in 2017. She is currently pursuing the Ph.D. degree with the Department of Mechanical, Industrial and Systems Engineering, The University of Rhode Island, USA. Her current research interests include adaptive neural networks and nonlinear systems, fault diagnosis, and deterministic learning theory.
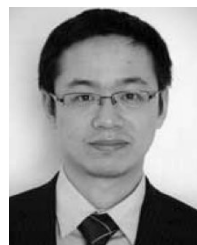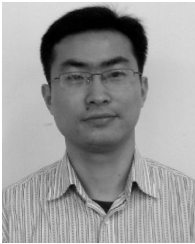
**QINGBIN GAO** (Member, IEEE) received the B.S. degree in mechanical engineering from the Harbin Institute of Technology, China, in 2011, and the Ph.D. degree in mechanical engineering from the University of Connecticut, in 2015. He has been an Assistant Professor with the Department of Mechanical and Aerospace Engineering, California State University Long Beach, from 2015 to 2018. From 2018 to 2019, he was an Assistant Professor with the Department of Mechanical Engineering, The University of Alabama. Since August 2019, he has been an Associate Professor with the School of Mechanical Engineering and Automation, Harbin Institute of Technology, Shenzhen. His main research focuses on the stability analysis and control synthesis of time-delay systems with applications to multiagent systems, manufacturing, connected vehicles, human learning, power systems, and vibrations. He was a recipient of the Best Conference Paper Award of the 19th International Conference on Networking, Sensing, and Control (ICNSC), in 2017, and the 6th American Society of Mechanical Engineers (ASME) Dynamic Systems and Control Conference (DSCC), in 2013.

**CHENGZHI YUAN** (Member, IEEE) received the B.S. and M.S. degrees in control theory and applications from the South China University of Technology, Guangzhou, China, in 2009 and 2012, respectively, and the Ph.D. degree in mechanical engineering from North Carolina State University, Raleigh, NC, USA, in 2016. He is currently an Assistant Professor with the Department of Mechanical, Industrial and Systems Engineering, The University of Rhode Island, Kingston, RI, USA. He has authored or coauthored over 90 journal articles and conference papers. His research interests span over general areas of dynamic systems and control theory, with particular focuses on analytical adaptive learning and control, hybrid systems, and multirobot distributed control. He has served extensively as an associate editor, the chair, the co-chair, and a program committee member in numerous international conferences. In particular, he served in the Technical Program Committee for the 2018 American Control Conference. He has served as a Guest Editor for special issue in the *Journal of IET Control Theory and Applications*, in 2018, and the *Journal of Advances in Mechanical Engineering*, in 2017.

**SHI-LU DAI** (Member, IEEE) received the B.Eng. degree in thermal engineering and the M.Eng. and Ph.D. degrees in control science and engineering from Northeastern University, Shenyang, China, in 2002, 2006, and 2010, respectively. From 2007 to 2009, he was a Visiting Student with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore. From 2015 to 2016, he was a Visiting Scholar with the Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN, USA. Since 2010, he has been with the School of Automation Science and Engineering, South China University of Technology, Guangzhou, China, where he is currently a Professor. His current research interests include adaptive and learning control and distributed cooperative systems.

**WEI ZENG** received the B.S. degree from the School of Mechanical and Electrical Engineering, South China University of Technology, Guangzhou, China, in 2002, the M.S. degree from the Department of Automation, Xiamen University, Xiamen, China, in 2008, and the Ph.D. degree from the College of Automation Science and Engineering, South China University of Technology, in 2012. From 2012 to 2014, he was a Post-Doctoral Fellow with the South China University of Technology. From 2017 to 2018, he was a Visiting Scholar with the Faculty of Health Sciences, The University of Sydney, Australia. He is currently a Professor with Longyan University, Longyan, China. His current research interests include biomechanics, biomedical signal processing, neural network-based nonlinear system identification, gait analysis, and dynamical pattern recognition.

**CONG WANG** (Member, IEEE) received the B.E. and M.E. degrees from the Beijing University of Aeronautics and Astronautics, in 1989 and 1997, respectively, and the Ph.D. degree from the Department of Electrical and Computer Engineering, National University of Singapore, in 2002. From 2001 to 2004, he did his postdoctoral research with the Department of Electronic Engineering, City University of Hong Kong. His research interests include intelligent control, neural networks, nonlinear systems and control, dynamical pattern recognition, pattern-based control, dynamical systems, and oscillation fault diagnosis.

• • •