Similar Fault Isolation of Discrete-Time Nonlinear Uncertain Systems Using Smallest Residual Principle

Jingting Zhang, Chengzhi Yuan, Paolo Stegagno

***Abstract*— This paper investigates the problem of similar fault isolation (sFI) for discrete-time nonlinear uncertain systems. The main challenge lies in that the differences among the so-called "similar" faults could be very small and easily hidden in system uncertainties. To overcome such a challenge, in this paper, the uncertain fault-induced system dynamics is first accurately identified using radial basis function neural network (RBF NNs), where the obtained knowledge can be stored and represented by constant RBF NNs. With the obtained constant networks, a bank of novel fault residual systems are designed by using an absolute measurement of fault dynamics difference, which can effectively measure the match level of the occurred fault from each trained fault. Based on the designed residual systems, real-time fault isolation decision making is achieved according to the smallest residual principle (SRP), i.e., the occurred fault is identified similar to one trained fault when the related residual becomes the smallest one among all the others. Rigorous analysis of the isolatability condition is also given. Extensive simulations have been conducted to demonstrate the effectiveness and advantages of the proposed approach.**

I. INTRODUCTION

In recent years, there have been many research activities in the development of fault isolation (FI) in various modern engineering systems (e.g., [1], [2], [3], [4], [5], [6]). Particularly, the model-based FI approach with the capabilities of providing a deep insight into the system process behavior attracts quite strong interests. For example, [7] designed an integrated robust FI architecture for distributed parameter systems. [8] constructed an entropy optimization filter for the FI problem of nonlinear non-Gaussian systems. In [9], a fast adaptive fault estimation (FAFE) approximator was proposed for linear systems. However, all these techniques virtually focus on systems with precisely known dynamics. The FI problem for nonlinear systems with unstructured uncertain dynamics remains as a rather challenging problem, especially when isolation of "similar" faults is concerned.

The "similar" faults are typically referred to those faults with small mutual differences that could be hidden by system uncertainties. For similar fault isolation (sFI) problem, over the past decades, lots of interesting techniques have been proposed but gained vry limited sucess, e.g., [10], [11], [12]. In [11], the FI problem for nonlinear uncertain systems was addressed by estimating the so-called fault mismatch functions. However, they have not achieved accurate modeling for system uncertain dynamics, thus are applicable only to those faults with the fault mismatch function in relatively large magnitudes (larger than that of the system uncertainty). [12] proposed a smallest residual principle (SRP) based FI scheme, where residuals are generated by measuring certain accumulated difference between occurred fault and each trained fault. With such schemes, if fault differences have frequently-changing signs, the associated residuals will approach zero, resulting in an isolation misjudgment.

For sFI problem of nonlinear uncertain systems, the difficulty lies in how to distinguish the fault difference from unmodeled system dynamics. Adaptive neural network (NN) provides a powerful tool. In [13], online approximation-based FI methods are developed using adaptive NNs to estimate the fault mismatch function. Nevertheless, since they fail to enable the NN weights converge to their optimal values without satisfying persistently exciting (PE) condition, the accurate modeling of fault mismatch functions cannot be achieved. To solve this problem, the deterministic learning (DL) theory was recently proposed and developed in [14], [15], [16], [17], [18], where a partial PE condition is guaranteed by a recurrent input to the radial basis function neural networks (RBF NNs), and then locally accurate modeling of the unknown system dynamics can be achieved, represented and stored as a constant RBF NN model. A DL-based sFI approach has been preliminarily investigated in [12] for nonlinear uncertain systems, which, however, has not addressed several important technical issues, such as how to expand applicability of the FI scheme to more stringent faults (e.g., faults with associated fault difference having frequently changing signs), and how to extend the methodology from continuous-time to discrete-time nonlinear uncertain systems.

In this paper, we investigate a new DL-based sFI approach for discrete-time nonlinear uncertain systems. First, motivated by [15], we propose an adaptive dynamics learning approach via DL to achieve a locally-accurate approximation of the unknown faulty system dynamics, enabling to distinguish the occurred similar faults from system uncertainties. The obtained knowledge is represented and stored in a constant RBF NN model [19]. Then, with the obtained knowledge, the residual systems are designed using absolute measurements of the faulty dynamics difference, to effectively evaluate the match level of the occurred fault to each trained fault. The FI decision making is based on SRP, i.e., the occurred fault is identified similar to one trained fault when the related residual becomes the smallest one among all the others.

This work was supported in part by the National Science Foundation under Grant CMMI 1929729, and in part by the Division of Research and Economic Development at University of Rhode Island under the Proposal Development Grant.

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)

P. Stegagno is with the Department of Electrical, Computer and Biomedical Engineering, University of Rhode Island, Kingston, RI 02881, USA (e-mail: pstegagno@uri.edu)

Rigorous analysis on isolatability condition is also given.

The contributions of this paper are: 1) the sFI problem for nonlinear uncertain systems is addressed, where the mutual differences among considered faults are so small that could be hidden by system uncertainties; 2) the sFI residual systems are developed based on absolute measurements of the faulty dynamics difference, to effectively evaluate the match level of the occurred fault to each trained fault.

The rest of this paper is organized as follows. Sec. II provides preliminary results and problem formulation. Sec. III presents the DL-based learning approach, the sFI scheme and the fault isolatability analysis. Simulation results are presented in Sec. IV. The conclusion is in Sec. V.

Notations. $\mathbb{R}$, $\mathbb{R}_+$ 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; $|\cdot|$ 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. $||x(k)||_1 = \frac{1}{K}\sum_{h=k-K}^{k-1}|x(h)|$ $(k \geq K > 1)$.

II. PRELIMINARIES AND PROBLEM FORMULATION

A. Preliminaries

The RBF NN is described by $f_{nn}(Z) = \sum_{i=1}^{Nn} w_i s_i(Z) = W^T S(Z)$[20], 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 $S(Z) = [s_1(||Z-\epsilon_1||), \cdots, s_{N_n}(||Z-\epsilon_{N_n}||)]^T$, with $s_i(\cdot)$ being a radial basis function, and ϵ_i $(i = 1, 2, \cdots, N_n)$ being distinct points in state space. The Gaussian function $s_i(||Z-\epsilon_i||) = exp[\frac{-(Z-\epsilon_i)^T(Z-\epsilon_i)}{\eta_i^2}]$ is one of the most commonly used RBF, with $\epsilon_i = [\epsilon_{i1}, \epsilon_{i2}, \cdots, \epsilon_{iq}]^T$ being the center of receptive field and η_i is the width of the receptive field. The Gaussian function belongs to the class of localized RBFs in the sense that $s_i(||Z-\epsilon_i||) \to 0$ as $||Z|| \to \infty$. Obviously, $S(Z)$ is bounded and there exists a real constant $S_M \in \mathbb{R}_+$ such that $||S(Z)|| \leqslant S_M$ [19].

As shown in [20], [21], 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 with sufficiently-large node number N_n, 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. $W^* \triangleq arg\min_{W\in\mathbb{R}^{N_n}}\{\sup_{Z\in\Omega_Z}|f(Z)-W^T S(Z)|\}$. Moreover, based on the localization property of RBF NNs [19], for any bounded trajectory $Z(t)$ within the compact set Ω_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 ϵ_ζ 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 $Z_p \in Z(k)$.

As shown in [19], 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 can be formally summarized in the following lemma.

Lemma 1 ([22]): 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 Ω_Z), the regressor subvector $S_\zeta(Z)$ consisting of RBFs with centers located in a small neighborhood of $Z(k)$ is PE.

B. Problem Formulation

Consider the following discrete-time nonlinear system:

$$\begin{aligned} x(k+1) =& f(x(k), u(k)) + v(x(k), u(k)) \\ &+ \beta(k-k_0)\phi^s(x(k), u(k)), \end{aligned} \tag{1}$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$ are system state and input; $f(x, u): \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is known nominal dynamics, $v(x, u): \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is modeling uncertainty, $\phi^s(x, u): \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is system dynamics deviation due to fault s $(s = 1, \cdots, N)$; $\beta(k-k_0)$ represents fault time profile: $\beta(k-k_0) = 0$ for $k < k_0$ and $\beta(k-k_0) = 1$ for $k \geq k_0$, k_0 is fault occurrence time. System (1) is said to operate in normal mode when $0 < k < k_0$ and s-th faulty mode when $k \geq k_0$. As typically adopted in the literature (e.g., [19]), we assume that the system trajectories in normal mode, i.e., (x^0, u^0), and all possible faulty modes, i.e., (x^s, u^s), are recurrent.

Following similar definition in [12], the "similar" faults are characterized as: 1) their mutual differences are small and could be hidden by system uncertainty, i.e., the magnitude of each $\phi^s(x, u) - \phi^{\bar{s}}(x, u)$ (with $s \in \{1, \cdots, N\}$, $\bar{s} \in \{1, \cdots, N\}/\{s\}$) could be smaller than that of system uncertain dynamics $v(x, u)$, and 2) the s-th faulty system trajectory (x^s, u^s) is close to each $\bar{s}$-th one $(x^{\bar{s}}, u^{\bar{s}})$, i.e.,

$$\begin{aligned} dist((x^s, u^s), (x^{\bar{s}}, u^{\bar{s}})) :=& \max\{\min ||(x^s, u^s), (x^{\bar{s}}, u^{\bar{s}})||\} \\ <& d_\zeta \end{aligned} \tag{2}$$

with a constant $0 < d_\zeta < d_\zeta^*$, and d_ζ^* is the size of the NN approximation region to be given later. We aim to achieve sFI for the system (1) by using real-time information of (x, u).

III. MAIN RESULTS

A. Fault Dynamics Identification

Consider the system (1) operating in s-th faulty mode, i.e.,

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

where $s = 0, 1, \cdots, N$ with $s = 0$ representing the normal mode (i.e., $\phi^0(x, u) = 0$). Note that the terms $v(x, u)$ and $\phi^s(x, u)$ cannot be decoupled from each other, we consider them together and define a general fault function $\eta^s(x, u) := v(x, u) + \phi^s(x, u)$. Then, Eq. (3) can be rewritten as:

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

[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 [19].

The goal here is to accurately identify the function $\eta^s(x, u)$.

According to Sec. II-A, for the function $\eta^s(x, u)$, there exists an ideal constant NN weight $W^{s*} \in \mathbb{R}^{N_n \times n}$ (with N_n denoting the number of NN nodes) such that

$$\eta_i^s(x, u) = W_i^{s*T} S(x, u) + \epsilon_{i,0}^s, \quad i = 1, \cdots, n, \tag{5}$$

where $S(x, u) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^{N_n}$ is a smooth RBF vector and $\epsilon_{i,0}^s$ is the estimation error satisfying $|\epsilon_{i,0}^s| < \epsilon^*$ with ϵ^* being a positive constant that can be made arbitrarily small given sufficiently large number of neurons.

Then, using the Gaussian RBF networks, for system (4), we construct the following dynamical identifier:

$$\begin{aligned}\hat{x}_i(k+1) =& a_i(\hat{x}_i(k) - x_i(k)) + f_i(x(k), u(k)) \\ &+ \hat{W}_i^{sT}(k+1) S(x(k), u(k)),\end{aligned} \tag{6}$$

where $\hat{x}_i \in \mathbb{R}$ is the i-th estimator state, x_i is the i-th state in (4), $0 < |a_i| < 1$ is a design parameter, and $\hat{W}_i^s \in \mathbb{R}^{N_n}$ is the estimate of W_i^{s*} following the adaption law:

$$\begin{aligned}&\hat{W}_i^s(k+1) = \hat{W}_i^s(k) \\ &- \frac{c_i(\tilde{x}_i(k) - a_i \tilde{x}_i(k-1)) S(x(k-1), u(k-1))}{1 + S^T(x(k-1), u(k-1)) S(x(k-1), u(k-1))}\end{aligned} \tag{7}$$

with design constant $0 < c_i < 2$ and $\tilde{x}_i(k) := \hat{x}_i(k) - x_i(k)$.

With the identifier (6) – (7), given the recurrent s-th faulty system trajectory (x^s, u^s) of (4), the partial PE condition of $S(x^s, u^s)$ is satisfied according to Lem. 1. From [15], the RBF NN weights $\hat{W}_i^s$ will converge to a small neighborhood of W_i^{s*} along the recurrent trajectory (x^s, u^s), and thus a locally accurate approximation of $\eta_i^s(x^s, u^s)$ in (4) can be achieved by the RBF NNs $\hat{W}_i^{sT} S(x^s, u^s)$ along the trajectory (x^s, u^s) after the transient process of convergence, i.e.,

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

with estimation error $|\epsilon_{i,1}^s| = O(\epsilon^*)$. Thanks to the convergence of $\hat{W}_i^s$, from [15], a locally accurate approximation of $\eta_i^s(x^s, u^s)$ is achieved by a constant model $\bar{W}_i^{sT} S(x^s, u^s)$:

$$\eta_i^s(x^s, u^s) = \bar{W}_i^{sT} S(x^s, u^s) + \epsilon_{i,2}^s, \tag{9}$$

where $\bar{W}_i^s := \frac{1}{K_2} \sum_{k=K_1}^{K_1+K_2-1} \hat{W}_i^s(k)$ with $[K_1, K_1+K_2-1]$ representing a time segment after the transient process, and $\epsilon_{i,2}^s$ is estimation error satisfying $|\epsilon_{i,2}^s| = O(\epsilon^*)$.

As argued in [15], the constant RBF NN $\bar{W}_i^{sT} S(x^s, u^s)$ has a certain ability of generalization, in the sense that its locally-accurate approximation for $\eta_i^s(x^s, u^s)$ is achieved in a local region Ω_ζ^s along the trajectory (x^s, u^s), that is,

$$\eta_i^s(x, u) = \bar{W}_i^{sT} S(x, u) + \epsilon_i^s, \quad \forall (x, u) \in \Omega_\zeta^s, \tag{10}$$

where $|\epsilon_i^s| \le \varepsilon^*$, $\Omega_\zeta^s := \{(x, u) \mid dist\Big((x, u), (x^s, u^s)\Big) < d_\zeta^*\}$ with $d_\zeta^* > 0$ characterizing the size of NN approximation region, and $\varepsilon^* = O(\epsilon^*)$ is approximation error within Ω_ζ^s.

According to the similar fault assumption specified in (2), for each s ($s \neq 0$, $s \in \{1, 2, \cdots, N\}$) and $\bar{s}$ ($\bar{s} \neq 0$, $\bar{s} \in \{1, 2, \cdots, N\}/\{s\}$), the faulty trajectory $(x^{\bar{s}}, u^{\bar{s}})$ satisfies: $dist((x^s, u^s), (x^{\bar{s}}, u^{\bar{s}})) < d_\zeta^*$, implying $(x^{\bar{s}}, u^{\bar{s}}) \in \Omega_\zeta^s$. From (10), once the system (1) operates in $\bar{s}$-th faulty mode, i.e., $(x, u) = (x^{\bar{s}}, u^{\bar{s}}) \in \Omega_\zeta^s$, we have $\eta_i^s(x^{\bar{s}}, u^{\bar{s}}) = \bar{W}_i^{sT} S(x^{\bar{s}}, u^{\bar{s}}) + \epsilon_i^s$ (with $|\epsilon_i^s| \le \varepsilon^*$) holds. However, if the system (1) operates in 0-th faulty mode (i.e., normal mode), the trajectory $(x, u) = (x^0, u^0)$ may not be sufficiently close to the faulty trajectory (x^s, u^s), i.e., $(x^0, u^0) \notin \Omega_\zeta^s$, such that $|\bar{W}_i^{sT} S(x^0, u^0) - \eta_i^s(x^0, u^0)| \not\le \varepsilon^*$. Similarly, if the system (1) operates in s-th faulty mode, the trajectory $(x, u) = (x^s, u^s) \notin \Omega_\zeta^0$, $|\bar{W}_i^{0T} S(x^s, u^s) - \eta_i^0(x^s, u^s)| \not\le \varepsilon^*$. Based on this, we make the following assumption:

Assumption 1: Consider the system trajectory (x, u) generated from (1), the constant RBF NN $\bar{W}_i^{sT} S(x, u)$ ($i = 1, \cdots, n$, $s \in \{0, 1, \cdots, N\}$) and the local region Ω_ζ^s from (10). Even if $(x, u) \notin \Omega_\zeta^s$, there exists a known constant $\gamma_i > 0$ such that $|\bar{W}_i^{sT} S(x, u) - \eta_i^s(x, u)| \le \gamma_i$.

Remark 1: The constant γ_i is not necessarily a large number, since the system trajectory (x, u) cannot be far away from the region Ω_ζ^s no matter which mode the system (1) operates in. Particularly, if the size of Ω_ζ^s is made large enough via sufficient training [15], [23], such that the system trajectory (x, u) keeps close to (or staying within) the region Ω_ζ^s, then γ_i will be relatively small.

B. Residual System Design and Isolation Scheme

Assuming that a fault l ($l \in \{1, \cdots, N\}$) occurs, we consider the following monitored system:

$$\begin{aligned}x(k+1) =& f(x(k), u(k)) + v(x(k), u(k)) \\ &+ \beta(k - k_0) \phi^l(x(k), u(k))\end{aligned} \tag{11}$$

with k_0 being the fault occurrence time. A bank of residual systems embedded with a novel mechanism of absolute measures of faulty dynamics difference are proposed:

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

where $e_i^s(0) = 0$, $i = 1, \cdots, n$, $s = 0, 1, \cdots, N$, x_i is the i-th state of (11); $f_i(x, u)$ is known dynamics in (11); $0 \le b_i < 1$ is a design parameter; $\bar{W}_i^{sT} S(x, u)$ is a constant RBF NN in (10) to represent $\eta_i^s(x, u) = v_i(x, u) + \phi_i^s(x, u)$.

Remark 2: In (10), $\bar{W}_i^{sT} S(x(k-1), u(k-1)) + f_i(x(k-1), u(k-1))$ in (12) is able to approximate the s-th trained dynamics $x_i^s(k) = f_i(x(k-1), u(k-1)) + v_i(x(k-1), u(k-1)) + \phi_i^s(x(k-1), u(k-1))$, and thus $|\bar{W}_i^{sT} S(x(k-1), u(k-1)) + f_i(x(k-1), u(k-1)) - x_i(k)|$ in (12) essentially measures the absolute difference between the monitored system dynamics in (11) and the s-th trained faulty dynamics.

With the above residual systems (12), a sFI scheme will be proposed based on the SRP [12]. Specifically, once a fault l occurs in the monitored system (11) (after time k_0), the residual e_i^l will decrease and become the smallest one among all the residuals obtained e_i^s ($s = 0, 1, \cdots, N$) from (12), enabling the occurred fault to be isolated.

Fault isolation decision scheme: Consider the residual systems (12). If there exists $l \in \{1, \cdots, N\}$, and a finite time $k_l > 0$ such that for all $i = 1, \cdots, n$ and $\bar{l} \in \{0, 1, \cdots, N\}/\{l\}$, $e_i^l(k_l) < e_i^{\bar{l}}(k_l)$, then the occurred fault will be identified similar to fault l at time k_l.

C. Fault Isolatability Condition

When a fault l occurs in the monitored system (11), we have for $k \geq k_0$, $\beta(k-k_0)\phi_i^l(x,u) = \phi_i^l(x,u)$ and $(x,u) = (x^l, u^l)$. By introducing a fault mismatch function: $\varrho_i^{s,l}(x,\, u) := \phi_i^s(x,\, u) - \phi_i^l(x,\, u)$ to represent the difference between the occurred fault l with each trained fault s, from (11), the residual system (12) can be rewritten as:

$$\begin{aligned} e_i^s(k) = b_i^{k-k_0} e_i^s(k_0) + \sum_{h=k_0}^{k-1} b_i^{k-1-h} |\bar{W}_i^{sT} S(x(h), u(h) \\ - \eta_i^s(x(h), u(h) + \varrho_i^{s,l}(x(h), u(h))|, \quad k > k_0. \end{aligned} \tag{13}$$

Note that the system trajectory $(x,u) = (x^l, u^l) \in \Omega_\zeta^s$ ($s \neq 0$, $s \in \{1, 2 \cdots, N\}$), we have $\eta_i^s(x^l, u^l) = \bar{W}_i^{sT} S(x^l, u^l) + \epsilon_i^s$ (with $|\epsilon_i^s| \leq \varepsilon^*$) from (10). Eq. (13) is obtained as:

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

It should be noted that since $(x,u) = (x^l, u^l) \notin \Omega_\zeta^0$, the function $\eta_i^0(x,u)$ cannot be described by (10) (as discussed above in the paragraph before Assumption 1), and the 0-th residual system cannot be rewritten as the form of (14).

Theorem 1 (Fault Isolatability): Consider the monitored system (11) and the residual system (12). If for $l \in \{1, 2, \cdots, N\}$, $\bar{l} \in \{0, 1, \cdots, N\}/\{l\}$, and all $i = 1, 2, \cdots, n$, there exists a constant $\mu_i > \frac{\varepsilon^*}{1-b_i}$ such that

$$\begin{aligned} \sum_{h=k_0}^{k-1} b_i^{k-1-h} |\varrho_i^{\bar{l},l}(x(h),\, u(h))| \geq \mu_i + \frac{\varepsilon^*}{1-b_i}, \\ \sum_{h=k_0}^{k-1} b_i^{k-1-h} |\phi_i^l(x(h),\, u(h))| \geq \mu_i + \frac{\gamma_i}{1-b_i}; \end{aligned} \tag{15}$$

then there must exist a finite time $k_l > k_0$ such that $e_i^l(k_l) < e_i^{\bar{l}}(k_l)$ holds, and the occurred fault is identified similar to the trained fault l at time k_l.

Proof: We first consider the 0-th residual system (13). Since $(x,u) = (x^l, u^l) \notin \Omega_\zeta^0$, we have $|\bar{W}_i^{0T} S(x,u) - \eta_i^0(x,u)| \leq \gamma_i$ from Ass. 1. Note that $e_i^0(k_0) \geq 0$, $\varrho_i^{0,l}(x,u) = -\phi_i^l(x,u)$ and from (15), we have

$$\begin{aligned} e_i^0(k) \geq \sum_{h=k_0}^{k-1} b_i^{k-1-h} (|\varrho_i^{0,l}(x(h), u(h))| - |\bar{W}_i^{0T} S(x(h), u(h) \\ - \eta_i^0(x(h), u(h)|) \geq \sum_{h=k_0}^{k-1} b_i^{k-1-h} (|\phi_i^l(x(h), u(h))| - \gamma_i) \\ > \sum_{h=k_0}^{k-1} b_i^{k-1-h} |\phi_i^l(x(h), u(h))| - \frac{\gamma_i}{1-b_i} \geq \mu_i. \end{aligned} \tag{16}$$

Consider the $\bar{l}$-th ($\bar{l} \neq 0$) residual system (14). Note that $e_i^{\bar{l}}(k_0) \geq 0$, $|\epsilon_i^{\bar{l}}| \leq \varepsilon^*$, and from (15), we have

$$\begin{aligned} e_i^{\bar{l}}(k) \geq \sum_{h=k_0}^{k-1} b_i^{k-1-h} (|\varrho_i^{\bar{l},l}(x(h),\, u(h))| - |\epsilon_i^{\bar{l}}|) \\ \geq \sum_{h=k_0}^{k-1} b_i^{k-1-h} (|\varrho_i^{\bar{l},l}(x(h),\, u(h))| - \varepsilon^*) \\ > \sum_{h=k_0}^{k-1} b_i^{k-1-h} |\varrho_i^{\bar{l},l}(x(h),\, u(h))| - \frac{\varepsilon^*}{1-b_i} \geq \mu_i. \end{aligned} \tag{17}$$

Moreover, consider the l-th residual system (14). Note that $\varrho_i^{l,l}(x,\, u) = 0$, and $|\epsilon_i^l| \leq \varepsilon^*$, we obtain

$$\begin{aligned} e_i^l(k) = b_i^{k-k_0} e_i^l(k_0) + \sum_{h=k_0}^{k-1} b_i^{k-1-h} |\epsilon_i^l| \leq \\ b_i^{k-k_0} e_i^l(k_0) + \sum_{h=k_0}^{k-1} b_i^{k-1-h} \varepsilon^* < b_i^{k-k_0} e_i^l(k_0) + \frac{\varepsilon^*}{1-b_i}. \end{aligned} \tag{18}$$

From (16) – (18), $e_i^l(k) < e_i^{\bar{l}}(k)$ is guaranteed for $\forall\, \bar{l} \in \{0, 1, \cdots, N\}/\{l\}$ if

$$b_i^{k-k_0} e_i^l(k_0) + \frac{\varepsilon^*}{1-b_i} \leq \mu_i. \tag{19}$$

Note that $\mu_i > \frac{\varepsilon^*}{1-b_i}$, and the right-hand side of the above inequality is a decreasing function of time step k, implying there exists a finite time $k_l > k_0$ such that the inequality (18) is satisfied, i.e., $e_i^l(k_l) < e_i^{\bar{l}}(k_l)$ holds and the occurred fault l can be isolated at time k_l. This ends the proof. ■

Remark 3: In (15), the first constraint on $\varrho_i^{\bar{l},l}(x,u)$ is not stringent, because the lower bound $\mu_i + \frac{\varepsilon^*}{1-b_i}$ (with $\mu_i > \frac{\varepsilon^*}{1-b_i}$) is basically depending on ε^* (be arbitrarily small by constructing sufficiently large number of neurons in the identification phase) and b_i (be freely selected from $0 \leq b_i < 1$). Thus, even when the the difference between fault l and fault $\bar{l}$ is relatively small (i.e., the occurred fault l is quite similar to mismatched fault $\bar{l}$), the condition (15) is satisfied and the occurred fault can be isolated. On the other hand, the constraints (15) are also loose owing to the small lower bound $\mu_i + \frac{\gamma_i}{1-b_i}$, where the constant γ_i can be made small following the arguments in Remark 1.

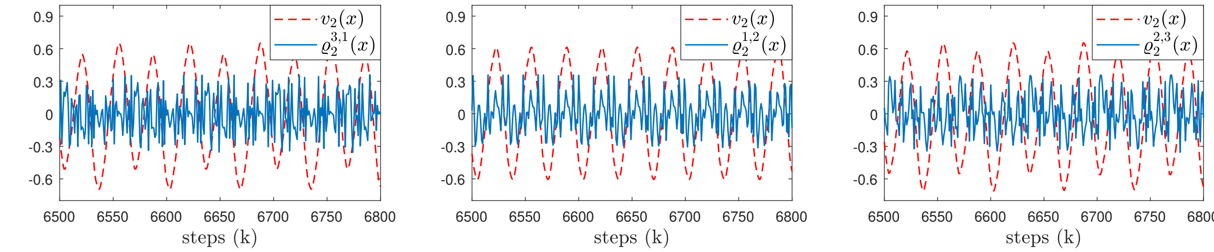
IV. SIMULATION STUDIES

Consider the Duffing oscillator system [19], through Euler approximation method, we have:

$$\begin{aligned} x_1(k+1) &= x_1(k) + T_s x_2(k), \\ x_2(k+1) &= x_2(k) + T_s(-p_2 x_1(k) - p_3 x_1^3(k) \\ &- p_1 x_2(k) + p_4 \cos(wT_s k)) + \beta(k-k_0)\phi_2^s(x(k)), \end{aligned} \tag{20}$$

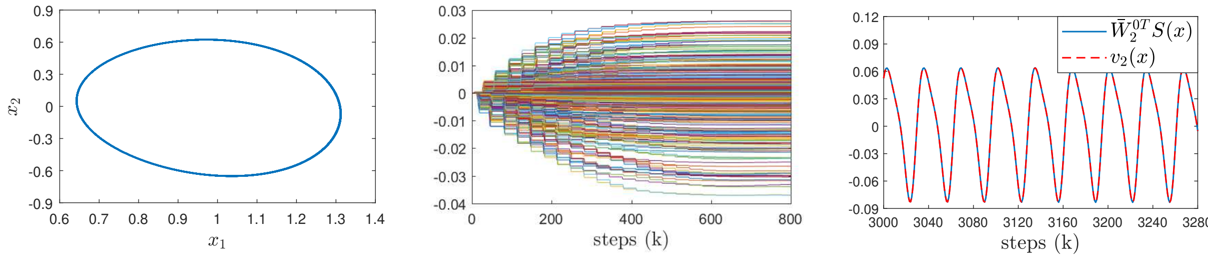
with states x_1, x_2, parameters $p_1 = 0.55$, $p_2 = -1.1$, $p_3 = 1$, $p_4 = 0.62$, $w = 1.9$, sampling period $T_s = 0.1s$; fault occurrence time $k_0 = 300$, fault functions $\phi_2^s(x)$ ($s = 1, 2, 3$) with $\phi_2^1(x(k)) := 0.18T_s \cos(20x_2(k) + 10x_1(k))$ denoting fault 1, $\phi_2^2(x(k)) := 0.18T_s \sin(10x_2(k) + 10x_1(k))$ fault 2, and $\phi_2^3(x(k)) := 0.18T_s \cos(30x_2(k) + 10x_1(k))$ fault 3. In this example, we assume that $f_1(x(k)) := x_1(k) + T_s x_2(k)$, $f_2(x(k)) := x_2(k) + T_s p_4 \cos(wT_s k)$ are the known nominal dynamics of the system, and $v_2(x(k)) := -T_s p_2 x_1(k) - T_s p_3 x_1^3(k) - T_s p_1 x_2(k)$ is the system uncertainty. Fig. 1 illustrates the comparison between fault mismatch functions $\varrho_2^{s,l}(x) = \phi_2^s(x) - \phi_2^l(x)$ ($s, l \in \{1, 2, 3\}$, $s \neq l$) and system uncertainty $v_2(x)$, showing that the difference between the fault s and fault l is relatively small and can be easily hidden by the system uncertainty. It is indicated that these three types of faults are "similar" faults that are difficult to isolate.

In the identification phase, for each s-th faulty mode ($s = 0, 1, 2, 3$, with $s = 0$ representing the normal mode, i.e., $\phi_2^0(x) = 0$), we employ the proposed identifier consisting



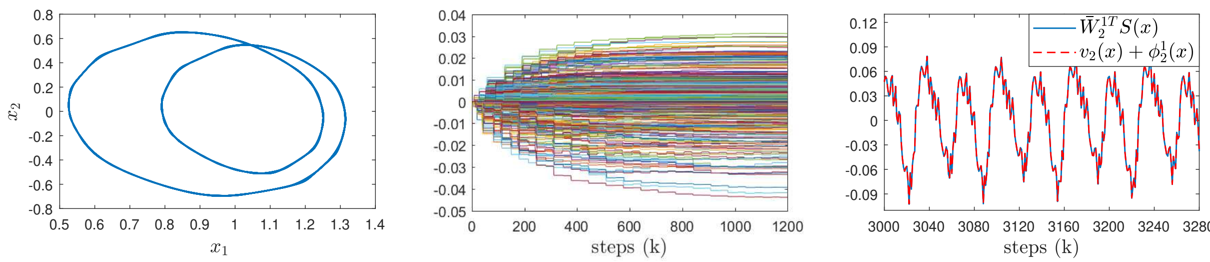
(a) $v_2(x)$ and $\varrho_2^{3,1}(x)$. (b) $v_2(x)$ and $\varrho_2^{1,2}(x)$. (c) $v_2(x)$ and $\varrho_2^{2,3}(x)$.

Fig. 1: Time profiles of system uncertainties $v_2(x)$ and the fault mismatch functions $\varrho_2^{s,l}(x)$ ($s, l \in \{1, 2, 3\}$, $s \neq l$).



(a) Normal system trajectory. (b) Weight convergence of $\hat{W}_2^0$. (c) Function approximation.

Fig. 2: Identification performance of the system uncertain dynamics $v_2(x)$ in normal mode.



(a) System trajectory in faulty mode 1. (b) Weight convergence of $\hat{W}_2^1$. (c) Function approximation.

Fig. 3: Identification performance of the system uncertain dynamics $v_2(x) + \phi_2^1(x)$ in faulty mode 1.

of (6) and (7) to accurately identify the uncertain dynamics $v_2(x) + \phi_2^s(x)$. Specifically, we construct the RBF networks $\hat{W}_2^{sT}S(x)$ in a regular lattice, with nodes $N_n = 29 \times 49$, the centers evenly spaced on $[0.4, 1.8] \times [-1.2, 1.2]$ and the widths $\eta_t = 0.05$ ($t = 1, 2, \cdots, 1421$). The design parameters for (6) and (7) are $a_2 = 0.1$ and $c_2 = 0.4$. The initial conditions are set as $\hat{W}_2^s(0) = 0$ and $x(0) = [0.2,\ 0.4]^T$. Consider the normal mode (faulty mode 0) of system (20), Fig. 2a shows that the normal system trajectory is recurrent. Fig. 2b implies the convergence of the weights $\hat{W}_2^0$ (each curve represents one weight parameter), based on which the constant weights $\bar{W}_2^0$ can be further obtained by $\bar{W}_2^0 = \frac{1}{100}\sum_{k=701}^{800}\hat{W}_2^0(k)$. Fig. 2c demonstrates that the obtained constant network $\bar{W}_2^{0T}S(x)$ achieves accurate approximation of $v_2(x)$. Similarly, the learning performance for the faulty mode 1 of system (20) is plotted in Fig. 3. As for the other two faulty modes, the simulation results are similar to those of the normal mode (and faulty mode 1), and thus omitted here owing to the space limitation.

After achieving accurate identification of all the uncertain faulty dynamics $v_2(x) + \phi_2^s(x)$ ($s = 0, 1, 2, 3$), we construct the following residual systems:

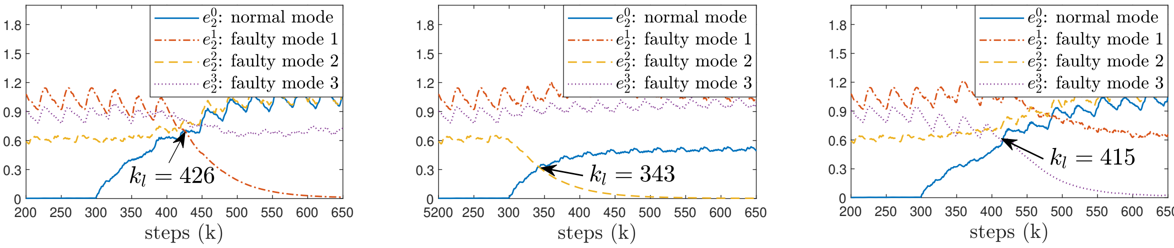
$$\begin{aligned} e_2^s(k) &= b_2 e_2^s(k-1) + |\bar{W}_2^{sT}S(x(k-1)) \\ &\quad + x_2(k-1) + T_s p_4 \cos(wT_s(k-1)) - x_2(k)|, \end{aligned} \tag{21}$$

with setting $b_2 = 0.98$. Consider fault 1 occurs in the monitored system (20) at time $k_0 = 300$. The real-time residuals generated from (21) are plotted in Fig. 4a, where the fault 1 is isolated at $k_l = 426$. Similar results of the cases when fault 2 or fault 3 occurs can also be observed in Figs. 4b and 4c, showing fault 2 and fault 3 are isolated at time $k_l = 343$ and $k_l = 415$, respectively.

To compare our scheme with the existing SRP-based FI scheme [12], we implement the estimators in [12] as follows:

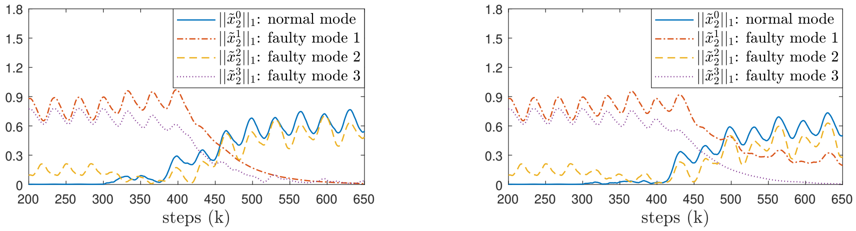
$$\begin{aligned} \hat{x}_2^s(k) &= b_2(\hat{x}_2^s(k-1) - x_2(k-1)) + x_2(k-1) \\ &\quad + T_s p_4 \cos(wT_s(k-1)) + \bar{W}_2^{sT}S(x(k-1)), \end{aligned} \tag{22}$$

where $b_2 = 0.98$ and $\bar{W}_2^{sT}S(x)$ ($s = 0, 1, 2, 3$) are exactly the above-obtained identification results. The residual signals for FI decision making are generated using an average L_1 norm, i.e., $||\tilde{x}_2^s(k)||_1 = \frac{1}{K}\sum_{h=k-K}^{k-1}|\hat{x}_2^s(h) - x_2(h)|$ with $K = 10$. From [12], when fault 1 occurs, the generated residuals $||\tilde{x}_2^s(k)||_1$ are plotted in Fig. 5a. It is shown that after the fault 1 occurs at time $k = 300$, both residuals $||\tilde{x}_2^1(k)||_1$ and $||\tilde{x}_2^3(k)||_1$ decrease to zero, and cannot be distinguished from each other. One explanation for such a phenomenon is that the fault difference $\varrho_2^{3,1}(x)$ has frequently-changing signs (as shown in Fig. 1a), its accumulated effect is easily offset, making the mismatched residual $||\tilde{x}_2^3||_1$ approach zero. A similar problem is also encountered in the process of FI for fault 3 (see Fig. 5b). Due to $\varrho_2^{2,3}(x)$ frequently changing sign,



(a) Isolation of fault 1. (b) Isolation of fault 2. (c) Isolation of fault 3.

Fig. 4: Fault isolation results using the proposed sFI scheme.



(a) Isolation of fault 1. (b) Isolation of fault 3.

Fig. 5: Fault isolation results using the existing sFI scheme in [12].

the mismatched residual $||\tilde{x}_2^2||_1$ keeps close to zero and even becomes the smallest in the time interval $[400, 480]$, leading to a FI misjudgment. All such issues can be overcome by our designed residual systems (21) (see Figs. 4a and 4c).

V. CONCLUSIONS

In this paper, we have proposed a novel adaptive dynamics learning-based sFI scheme for discrete-time nonlinear uncertain systems. First, to deal with the system uncertainty, we achieved locally-accurate approximation for the system uncertain dynamics under both normal mode and all faulty modes, where the learned knowledge was obtained and stored in constant RBF networks. Then, a bank of novel residual systems were constructed by using the absolute measures of the fault dynamics difference. FI decision making was achieved based on the SRP. The effectiveness and advantages of the proposed sFI scheme have been verified through both rigorous analysis and numerical simulations. In future work, we plan to extend the results to explore applications of the proposed approaches to real engineering problems, such as manufacturing process and robotics.

REFERENCES

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

[2] J. Gertler, *Fault detection and diagnosis*. Springer, 2013.

[3] I. Hwang, S. Kim, Y. Kim, and C. E. Seah, "A survey of fault detection, isolation, and reconfiguration methods," *IEEE transactions on control systems technology*, vol. 18, no. 3, pp. 636–653, 2010.

[4] H. Chen and S. Lu, "Fault diagnosis digital method for power transistors in power converters of switched reluctance motors," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 2, pp. 749–763, 2013.

[5] Q. Zhao and Z. Xu, "Design of a novel knowledge-based fault detection and isolation scheme," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 34, no. 2, pp. 1089–1095, 2004.

[6] J. Seshadrinath, B. Singh, and B. K. Panigrahi, "Vibration analysis based interturn fault diagnosis in induction machines," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 1, pp. 340–350, 2014.

[7] 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.

[8] L. Guo, L. Yin, H. Wang, and T. Chai, "Entropy optimization filtering for fault isolation of nonlinear non-gaussian stochastic systems," *IEEE Transactions on Automatic Control*, vol. 54, no. 4, pp. 804–810, 2009.

[9] K. Zhang, B. Jiang, and V. Cocquempot, "Adaptive observer-based fast fault estimation," *International Journal of Control, Automation, and Systems*, vol. 6, no. 3, pp. 320–326, 2008.

[10] J. Zhang, C. Yuan, P. Stegagno, H. He, and C. Wang, "Small fault detection of discrete-time nonlinear uncertain systems," *IEEE transactions on cybernetics*, 2019, doi: 10.1109/TCYB.2019.2945629.

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

[12] 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.

[13] C. Keliris, M. M. Polycarpou, and T. Parisini, "A robust nonlinear observer-based approach for distributed fault detection of input–output interconnected systems," *Automatica*, vol. 53, pp. 408–415, 2015.

[14] C. Yuan and C. Wang, "Performance of deterministic learning in noisy environments," *Neurocomputing*, vol. 78, no. 1, pp. 72–82, 2012.

[15] ——, "Design and performance analysis of deterministic learning of sampled-data nonlinear systems," *Science China Information Sciences*, vol. 57, no. 3, pp. 1–18, 2014.

[16] ——, "Persistency of excitation and performance of deterministic learning," *Systems & control letters*, vol. 60, no. 12, pp. 952–959, 2011.

[17] C. Yuan, S. Licht, and H. He, "Formation learning control of multiple autonomous underwater vehicles with heterogeneous nonlinear uncertain dynamics," *IEEE Transactions on Cybernetics*, no. 99, pp. 1–15, 2017.

[18] W. Wu, C. Wang, and C. Yuan, "Deterministic learning from sampling data," *Neurocomputing*, vol. 358, pp. 456–466, 2019.

[19] C. Wang and D. J. Hill, *Deterministic learning theory for identification, recognition, and control*. CRC Press, 2009.

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

[21] M. POEWELL, "The theory of radial basis function approximation," 1992.

[22] C. Wang and D. J. Hill, "Learning from neural control," *IEEE Transactions on Neural Networks*, vol. 17, no. 1, pp. 130–146, 2006.

[23] W. Chen, S. Hua, and S. S. Ge, "Consensus-based distributed cooperative learning control for a group of discrete-time nonlinear multi-agent systems using neural networks," *Automatica*, vol. 50, no. 9, pp. 2254–2268, 2014.