Control Theory Tech, Vol. 14, No. 2, pp. 159–175, May 2016

Control Theory and Technology

http://link.springer.com/journal/11768

Fault detection for nonlinear discrete-time systems via deterministic learning

Junmin HU, Cong WANG†, Xunde DONG

School of Automation Science and Engineering, South China University of Technology, Guangzhou Guangdong 510640, China;

Guangdong Key Laboratory of Biomedical Engineering, Guangzhou Guangdong 510640, China

Received 21 October 2014; revised 25 March 2016; accepted 25 March 2016

Abstract

Recently, an approach for the rapid detection of small oscillation faults based on deterministic learning theory was proposed for continuous-time systems. In this paper, a fault detection scheme is proposed for a class of nonlinear discrete-time systems via deterministic learning. By using a discrete-time extension of deterministic learning algorithm, the general fault functions (i.e., the internal dynamics) underlying normal and fault modes of nonlinear discrete-time systems are locally-accurately approximated by discrete-time dynamical radial basis function (RBF) networks. Then, a bank of estimators with the obtained knowledge of system dynamics embedded is constructed, and a set of residuals are obtained and used to measure the differences between the dynamics of the monitored system and the dynamics of the trained systems. A fault detection decision scheme is presented according to the smallest residual principle, i.e., the occurrence of a fault can be detected in a discrete-time setting by comparing the magnitude of residuals. The fault detectability analysis is carried out and the upper bound of detection time is derived. A simulation example is given to illustrate the effectiveness of the proposed scheme.

Keywords: Fault detection, nonlinear discrete-time systems, deterministic learning, neural networks, locally accurate modeling

DOI 10.1007/s11768-016-4140-z

1 Introduction

The design and analysis of fault detection and isolation (FDI) for nonlinear systems are important issues in modern engineering systems. In complex nonlinear systems, the faults are often hidden among the modeling uncertainties. For nonlinear systems with structured modeling uncertainties, in which the faults generated from nonlinear structure and cannot be decoupled from the unknown inputs. For diagnosis of non-

†Corresponding author.

E-mail: wangcong@scut.edu.cn.

This work was supported by the National Science Fund for Distinguished Young Scholars (No. 61225014), the National Major Scientific Instruments Development Project (No. 61527811), the National Natural Science Foundation of China (Nos. 61304084, 61374119), the Guangdong Natural Science Foundation (No. 2014A030312005), and the Space Intelligent Control Key Laboratory of Science and Technology for National Defense.

linear systems with unstructured modeling uncertainties and nonlinear faults, the issue is much challenging since the faults cannot be decoupled from the unstructured modeling uncertainties. Over the past decades, much progress has been achieved for FDI of nonlinear continuous-time systems with structured and unstructured modeling uncertainties [1–14]. In particular, adaptive threshold approaches using neural networks have recently received much attention for FDI of nonlinear systems with unstructured modeling uncertainties [10–14]. However, only limited results has been obtained for fault diagnosis of nonlinear discrete-time systems [15–19]. On the other hand, due to the increasing popularity of applying digital computers in engineering, discrete-time systems can be more veritable to describe practical problem in control systems than continuous-time systems. The analysis and synthesis of fault detection (FD) for nonlinear discrete-time systems should be received more attention. Specifically, in [16], based on a parametric model of the faults, compensation of the fault effect on the state prediction is achieved via an adaptive discrete-time approach. A solution for fault isolation and identification is also proposed by using postfault analysis. In [18], an online approximation based fault detection and diagnosis scheme for multiple state or output faults was proposed for a class of nonlinear multiple-input-multiple-output (MIMO) discrete-time systems. The faults considered could be incipient or abrupt, and were modeled using input and output signals of the system. In [19], a distributed fault detection and isolation approach based on adaptive approximation was proposed for nonlinear uncertain large-scale discrete-time systems. Local and global FDI schemes were provided due to the utilization of specialized fault isolation estimators and a global fault diagnoser.

In adaptive approximation based fault detection and isolation of general nonlinear systems, however, convergence of the employed neural network (NN) weights to their optimal values and accurate neural network approximation of nonlinear fault functions were less investigated. Recently, a deterministic learning theory was proposed for identification, control and recognition of nonlinear dynamical systems exhibiting periodic or recurrent trajectories [20–22]. By using localized RBF neural networks, it is shown that almost any periodic or recurrent trajectory can lead to the satisfaction of a partial persistent excitation (PE) condition, which in turn yields accurate neural network approximation of the system dynamics in a local region along the periodic or recurrent trajectory. Further, rapid recognition of a test dynamical pattern from a set of training dynamical patterns is achieved by using the locally accurate neural network approximation of system dynamics [21, 22]. Based on deterministic learning theory, in [23] a FD scheme was proposed for rapid detection of small oscillation faults generated from nonlinear continuous-time systems. The modeling uncertainty and nonlinear fault functions are firstly accurately approximated via deterministic learning, and then the knowledge is utilized to achieve rapid detection of small oscillation faults.

In this paper, a fault detection scheme is proposed for a class of nonlinear discrete-time systems based on deterministic learning theory. First, the internal dynamics underlying normal and fault modes of nonlinear discrete-time systems are locally-accurately approximated by discrete-time dynamical RBF networks via a discrete-time extension of deterministic learning algorithm. Second, by utilizing the knowledge of system dynamics obtained through deterministic learning, a bank of estimators is constructed for the training normal and fault modes, and by comparing the set of estimators with the monitored system, a set of residuals are obtained and used to measure the differences between the dynamics of the monitored system and the dynamics of the trained systems. Finally, when the monitored system is in normal mode, the estimator corresponding to the normal mode will yield the smallest residual. When the monitored system is in one of the fault modes, the residual of the fault estimator corresponding to this fault mode will become small and the residual of the normal estimator will become large. This means that the occurrence of a fault can be detected in a discrete-time setting according to the smallest residual principle. The fault detectability analysis is carried out and the upper bound of detection time is derived. Compared with existing results on FD for nonlinear discrete-time systems, the main feature of this paper is that with locally-accurate approximation of the general fault functions of nonlinear discrete-time systems achieved via deterministic learning, the detection sensitivity to small faults of nonlinear discrete-time systems is increased substantially.

The rest of this paper is organized as follows. Section 2 begins with the introduction of deterministic learning theory, followed by the formulation of the problem in Section 3. In Section 4, the modeling and representation of the general fault function via deterministic learning for normal and fault modes is presented. On this basis, the detection scheme is given according to the smallest residual principle. Simulation results are included in Section 5. Section 6 concludes the paper.

2 Deterministic learning theory

In this section, we briefly review deterministic learning theory, which is developed for NN identification, recognition and control of nonlinear dynamical systems undergoing periodic or recurrent trajectory [21, 22]. In deterministic learning theory, identification of system dynamics of general nonlinear systems is achieved according to the following elements: i) employment of localized RBF networks; ii) satisfaction of a partial PE condition; iii) exponential stability of the adaptive system along the periodic or recurrent orbit; and iv) locally-accurate NN approximation of the unknown system dynamics.

The presented deterministic learning theory, especially the approach for identification and rapid recognition of dynamical patterns, provides a solution for the problem of rapid detection and isolation of oscillation faults generated from uncertain nonlinear systems [22, 23]. The result may be further applied to the analysis of practical problems. For instance, in [24], an approach for approximately accurate modeling and rapid detection of stall precursors based on deterministic learning and dynamical pattern recognition was developed. The studies on Mansoux model based on low-speed axial compressor of Beihang University were conducted to show the effectiveness of the approach. Zeng and Wang [25, 26] presented an algorithm to eliminate the effect of walking speed for efficient gait recognition in the lateral view. Then human gait dynamics underlying different individuals' gaits across different walking speeds were locally accurately approximated using deterministic learning. A method for electrocardiogram (ECG) pattern modeling and recognition via deterministic learning theory was also presented in [27].

2.1 RBF network and PE condition

The RBF networks can be described by

$$f_{nn}(Z) = \sum_{i=1}^{N} w_i s_i(Z) = W^{\mathrm{T}} S(Z), \tag{1}$$

where $Z \in \Omega_Z \subset \mathbb{R}^q$ is the input vector with q being the NN input dimension, $W = [w_1 \ \cdots \ w_N]^{\mathrm{T}} \in \mathbb{R}^N$ is the weight vector, N is the NN node number, and $S(Z) = [s_1(\|Z-\xi_1\|) \ \cdots \ s_N(\|Z-\xi_N\|)]^{\mathrm{T}}$, with $s_i(\cdot)$ being a radial basis function, and ξ_i $(i = 1,\dots,N)$ being distinct points in state space. The Gaussian function $s_i(\|Z-\xi_i\|) = \exp[\dfrac{-(Z-\xi_i)^{\mathrm{T}}(Z-\xi_i)}{\eta_i^2}]$ is one of the most commonly used RBF, where $\xi_i = [\xi_{i1} \ \cdots \ \xi_{iN}]^{\mathrm{T}}$ is the center of the receptive field and η_i is the width of the receptive field. The Gaussian function belongs to the class of localized radial basis functions in the sense that $s_i(\|Z-\xi_i\|) \to 0$ as $\|Z\| \to \infty$.

It has been shown in [20–22] that for any continuous function $f(Z) : \Omega_Z \to R$ where $\Omega_Z \subset \mathbb{R}^q$ is a compact set, and for the NN approximator, where the node number N is sufficiently large, there exists an ideal constant weight vector W^* ($W^* = [w_1^* \ \cdots \ w_N^*]^{\mathrm{T}} \in \mathbb{R}^N$), such that for each $\epsilon^* > 0$, $f(Z) = W^{*\mathrm{T}} S(Z) + \epsilon(Z)$, $\forall Z \in \Omega_Z$, where $|\epsilon(Z)| < \epsilon^*$ is the approximation error. Moreover, for any bounded trajectory $Z_\zeta(t)$ 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 φ_ζ:

$$f(Z) = W_\zeta^{*\mathrm{T}} S_\zeta(Z) + \epsilon_\zeta, \tag{2}$$

where $S_\zeta(Z) = [s_{j_1}(Z) \ \cdots \ s_{j_\zeta}(Z)]^{\mathrm{T}} \in \mathbb{R}^{N_\zeta}$, with $N_\zeta < N$, $|s_{j_i}| > \iota\,(j_i = j_1,\dots,j_\zeta)$, $\iota > 0$ is a small positive constant, $W_\zeta^* = [w_{j_1}^* \ \cdots \ w_{j_\zeta}^*]^{\mathrm{T}}$, and ϵ_ζ is the approximation error, with $\big||\epsilon_\zeta| - |\epsilon|\big|$ being small.

Based on previous results on the PE property of RBF networks [21], it is shown that for a localized RBF network $W^{\mathrm{T}} S(Z)$ whose centers are placed on a regular lattice, almost any recurrent trajectory $Z(t)$ can lead to the satisfaction of the PE condition of the regressor subvector $S_\zeta(Z)$ [22].

The following definition of persistency of excitation (PE) in discrete form provided by [28] is utilized.

Definition 1 A sequence $S(k) \in \mathbb{R}^n$ is said to be persistently exciting (in N steps), if there exists $N \in Z_+$, $\alpha > 0$ such that

$$\sum_{k=t_0}^{t_0+N-1} S(k) S^{\mathrm{T}}(k) \geqslant \alpha I \tag{3}$$

uniformly in t_0.

2.2 Deterministic learning of discrete-time systems

The deterministic learning theory is extended to modeling and control of nonlinear discrete-time systems in [29–31].

A nonlinear discrete-time system in the following form is considered:

$$X(k) = F(X(k-1),\dots,X(k-m);p), \tag{4}$$

where $X = [x_1 \ \cdots \ x_n]^{\mathrm{T}} \in \mathbb{R}^n$ is the state of the system, which is measurable, p is a system parameter vector, and different p will in general produce different dynamical patterns, $F(\cdot\,;p) = [f_1(\cdot\,;p)\ \cdots\ f_n(\cdot\,;p)]^{\mathrm{T}}$ is a smooth but unknown nonlinear vector field.

By using the following dynamical RBF network, spatially localized learning for discrete-time systems is achieved.

$$\begin{aligned}\hat{X}(k) &= -A(X(k-1) - Z(k-1)) + \hat{W}^{\mathrm{T}}(k)S_k, \\ Z(k) &= -A(X(k-1) - Z(k-1)) + \hat{W}^{\mathrm{T}}(k+1)S_k,\end{aligned} \tag{5}$$

where $\hat{X}(k) = [\hat{x}_1(k)\ \cdots\ \hat{x}_n(k)]^{\mathrm{T}} \in \mathbb{R}^n$ is the estimation of state vector, $X(k)$ is the state of system (4), $A = \mathrm{diag}\{a_1, \dots, a_n\}$ is diagonal matrix, $|a_i| < 1$ is the design constant, and $\hat{W}^{\mathrm{T}}(k)S_k$ is a RBF networks used to approximate the unknown nonlinearity $F(\cdot\,;p)$ of (4), $S_k = S(X(k-1), \dots, X(k-m))$, $\hat{W}(k) = [\hat{W}_1(k)\ \cdots\ \hat{W}_n(k)]$ is the weight estimate, and $\hat{W}_i$ is updated by

$$\hat{W}_i(k+1) = \hat{W}_i(k) + \Gamma_i S_k v_i(k), \tag{6}$$

where $\Gamma_i = \Gamma_i^{\mathrm{T}} > 0$, $v_i(k) = \dfrac{x_i(k) - \hat{x}_i(k)}{1 + S_k^{\mathrm{T}}\Gamma_i S_k}$.

The identification scheme [22] describes a method of adjusting the NN weights $\hat{W}$ in (6) in order to satisfy

$$\lim_{k\to\infty} \hat{W}(k) = W^*. \tag{7}$$

A constant vector of neural weights is chosen as

$$\bar{W} = \frac{1}{k_{\mathrm{b}} - k_{\mathrm{a}} + 1}\sum_{k=k_{\mathrm{a}}}^{k_{\mathrm{b}}} \hat{W}(k) \tag{8}$$

with $\{k_{\mathrm{a}}, \dots, k_{\mathrm{b}}\}$ representing a time segment after the transient process. We have that locally-accurate approximation of system dynamics along the tracking orbit φ_ζ can be obtained as follows [22]:

$$\begin{aligned}F(\cdot\,;p) &= W^{*\mathrm{T}}S(\varphi_\zeta) + \epsilon^* \\ &= \hat{W}^{\mathrm{T}}S(\varphi_\zeta) + \epsilon_1 = \bar{W}^{\mathrm{T}}S(\varphi_\zeta) + \epsilon_2,\end{aligned} \tag{9}$$

where ε^* is the ideal approximation error, ε_1 is the practical approximation error for using $\hat{W}^{\mathrm{T}}S(\varphi_\zeta)$, and ε_2 is the practical approximation error for using $\bar{W}^{\mathrm{T}}S(\varphi_\zeta)$. In deterministic learning, both convergence of partial neural network weights to their optimal values and locally-accurate approximation of system dynamics can be achieved. This implies that either $\hat{W}^{\mathrm{T}}S(\varphi_\zeta)$ or $\bar{W}^{\mathrm{T}}S(\varphi_\zeta)$ is indeed capable of approximating the system dynamics to the desired error level ε^*, i.e., $\varepsilon_1 = \mathrm{O}(\varepsilon^*)$ and $\varepsilon_2 = \mathrm{O}(\varepsilon_1) = \mathrm{O}(\varepsilon^*)$, which ensure that ε_1 and ε_2 are close to ε^*.

3 Problem formulation

Consider the following class of uncertain nonlinear discrete-time system:

$$\begin{aligned}x(k+1) =& f(x(k), u(k)) + \eta(x(k), u(k)) \\ &+ \beta(k-k_0)\phi^s(x(k), u(k)),\ \ k = 0, 1, \dots,\end{aligned} \tag{10}$$

where $x(k) = [x_1(k)\ \cdots\ x_n(k)]^{\mathrm{T}} \in \mathbb{R}^n$ is the state vector of the system, $u(k) = [u_1(k)\ \cdots\ u_m(k)]^{\mathrm{T}} \in \mathbb{R}^m$ is the control input vector, k is the discrete-time instant. $f(x(k), u(k)) = [f_1(x(k), u(k))\ \cdots\ f_n(x(k), u(k))]^{\mathrm{T}}$ is unknown smooth nonlinear vector field representing the dynamics of the normal model, $\eta(x(k), u(k))$ stands for the uncertainties including external disturbances, modeling errors and possibly discretization errors. The fault $\beta(k-k_0)\phi^s(x(k), u(k))$ is the deviation in system dynamics due to fault s $(s = 1, \dots, M)$, and $\beta(k-k_0)$ represents the fault time profile, with k_0 being the unknown fault occurrence time. When $k \leqslant k_0$, $\beta(k-k_0) = 0$, and when $k \geqslant k_0$, $\beta(k-k_0) = 1$.

The system state is assumed to be observable, and the system input is usually designed as a function of the system states. An assumption regarding the state and input of nonlinear discrete-time system is given as follows.

Assumption 1 The system states $x(k)$ are bounded for both normal and fault modes.

The state sequence $(x(k))_{k=0}^{\infty}$ of (10) with initial condition $x(0)$ is defined as the system trajectory. The trajectory in normal mode is denoted as $\varphi^0(x(0))$ or φ^0 for conciseness of presentation, and the trajectory in fault mode s is denoted as $\varphi^s(x(0))$ or φ^s for conciseness of presentation.

Assumption 2 The system trajectories for both normal and fault modes are recurrent trajectories.

Remark 1 A recurrent trajectory represents a large set of periodic and period-like trajectories generated from nonlinear dynamical systems, which includes periodic, quasi-periodic, almost-periodic, and even chaotic trajectories [32]. Roughly speaking, a recurrent trajectory is characterized as follows [33]: given a vector $\varsigma > 0$, there exists a $T(\varsigma) > 0$ such that the trajectory returns to the ς-neighborhood of any point on the trajectory within a time not greater than $T(\varsigma)$. A remarkable feature of the recurrent trajectories is that, regardless of the choice of the initial condition, for any given ς-neighborhood the whole trajectory lies in the -neighborhood of the segment of the trajectory corresponding to a bounded time interval $T(\varsigma)$.

Recurrent trajectories are common types of behaviors for nonlinear dynamical systems, which comprise the most important types (though not all types) of trajectories generated from nonlinear dynamical systems. Many practical dynamical systems, such as rotating machineries [34], electronic systems [35], power systems [36], [37], communication network [38], ECG systems [39], etc., can exhibit such kind of trajectories or oscillations. Therefore, this approach can be widely applied in the practical systems. Moreover, deterministic learning theory has been applied to the practical rotating machinery [24], gait silhouettes [25] and ECG systems [27].

Remark 2 In this paper, the system of fault detection is investigated with observable state. In practice, accurate measurement of all system states may not be available. Fault detection for the nonlinear systems with partial-state measurement can be studied by combining the approach proposed in this paper and the approach proposed in Chapter 7 of [22], in which deterministic learning algorithms with partial-state measurement are presented. Due to the limitation of space, this issue will not be pursued in this paper. In the practical systems, which exhibit periodic or recurrent trajectories, the state vectors are bounded (or oscillatory). For example, the currents of power distribution systems are oscillating in the normal situation. When high impedance faults occur, the oscillations will be distorted but still remain lower than the over current thresholds [37]. Such oscillating faults are similar to the normal behaviors and are very difficult to be detected.

In general, the uncertainties represented by the vector field $\eta(x(k), u(k))$ is unstructured. For the design and analysis of a detection scheme, the following assumption is needed.

Assumption 3 $\eta(x(k), u(k))$ is bounded in a compact region by some known function $\bar{\eta}_x(x(k), u(k))$, i.e., $\|\eta(x(k),\ u(k))\| \leqslant \bar{\eta}(x(k), u(k)), \forall (x(k), u(k)) \in \Omega \subset \mathbb{R}^n \times \mathbb{R}^m, \forall k \geqslant 0$.

When the system is in fault modes, since the fault functions may be hidden by the uncertainties, the uncertainties $\eta(x(k), u(k))$ and the fault functions $\beta(k - k_0)\phi^s(x(k), u(k))$ $(s = 1, \dots, M)$ cannot be decoupled from each other. The two terms are considered together as an undivided term, and is defined as the general fault function [23]:

$$\psi^s(x(k), u(k)) = \eta(x(k), u(k)) + \beta(k - k_0)\phi^s(x(k), u(k)), \tag{11}$$

where $s = 1, \dots, M$, $\phi^s(x(k), u(k))$ represents the sth fault belongs to the set of fault functions. For simplicity of presentation, the normal mode is represented by fault mode $s = 0$, with $\phi^0(x(k), u(k)) = 0$, i.e., $\psi^0(x(k), u(k)) = \eta(x(k), u(k))$.

For both normal and fault modes, combined with (11), system (10) can be represented by

$$x(k+1) = f(x(k), u(k)) + \psi^s(x(k), u(k)), \tag{12}$$

where $s = 0, 1, \dots, M$.

4 Fault detection for nonlinear discrete-time systems

4.1 Fault modeling and representation

Construct the following discrete-time dynamical RBF network:

$$\begin{cases} \hat{x}(k) = f(x(k), u(k)) + A(z(k-1) - x(k-1)) \\ \qquad\quad +\hat{W}^{s\mathrm{T}}(k)S(x(k)), \\ z(k) = f(x(k), u(k)) + A(z(k-1) - x(k-1)) \\ \qquad\quad +\hat{W}^{s\mathrm{T}}(k+1)S(x(k)), \end{cases} \tag{13}$$

where $A = \mathrm{diag}\{a_1, \dots, a_n\}$ is a diagonal matrix, with $0 < |a_i| < 1$ being design constant, $\hat{x}(k) = [\hat{x}_1(k)\ \hat{x}_2(k)\ \cdots\ \hat{x}_n(k)]^\mathrm{T}$ is the state vector of (13), $x(k) = [x_1(k)\ \cdots\ x_n(k)]^\mathrm{T}$ is the state vector of system (10), $z(k)$ is the prediction of $\hat{x}(k)$. The Gaussian RBF network $\hat{W}^{s\mathrm{T}}(k+1)S(x(k)) = [\hat{W}_1^{s\mathrm{T}}(k+1)S(x(k))\ \cdots\ \hat{W}_n^{s\mathrm{T}}(k+1)S(x(k))]$ $(s = 0, \dots, M)$ is used to approximate the general fault function (11).

To adjust the NN weights $\hat{W}$, the adaptive law (6) is designed as [30].

Remark 3 The parameters of learning system, including the adaptive law (6), constant vector of neural weights (8) and dynamical RBF network (13), are set according to the analytic results proposed in [22], [30] and [40]. For example, increasing the node number will decease the RBF approximation error, while the setting of the RBF widths will affect the level of persistent excitation, which in turn will also affect the learning capability of the proposed scheme. The observer gain a_i is small, increasing a_i will increase the convergence rate of the estimated parameters. When a_i increases to a certain extent, increasing a_i will decrease convergence rate of the estimated parameters and deteriorates the ability of the dynamical RBF network (13) to learn the dynamics of the fault. Please see [40] for the detail discussion of the relationship between a_i and the learning capability of the deterministic learning method. The way of setting Γ_i in the adaptive law (6) is similar to that of setting a_i.

Lemma 1 Consider the close-loop system consisting of the nonlinear discrete-time system (12), the dynamical RBF network (13) and the NN update law (6). For both normal and fault modes of (12), we have that all the signals in the closed-loop system remain bounded, and the general fault function $\psi^s(x(k), u(k))$ of system (12) is locally-accurately approximated along the trajectory φ^s by $\hat{W}^{\mathrm{T}}S(\varphi^s)$ as well as by $\bar{W}^{\mathrm{T}}S(\varphi^s)$ (see (8)).

Proof The proof of Lemma 1 can be found in [30].

Based on the convergence result of $\hat{W}$, we can obtain a constant vector of neural weights $\bar{W}$ according to (8), such that

$$\psi^s(x(k), u(k)) = \hat{W}^{s\mathrm{T}}S(\varphi^s) + \epsilon_1 = \bar{W}^{s\mathrm{T}}S(\varphi^s) + \epsilon_2, \quad (14)$$

where ϵ_1 and ϵ_2 are the approximation errors using $\hat{W}^{s\mathrm{T}}S(\varphi^s)$ and $\bar{W}^{s\mathrm{T}}S(\varphi^s)$, respectively. It is clear that after the transient process, $\big||\epsilon_1| - |\epsilon_2|\big|$ is small. In other word, the general fault function $\psi^s(x(k), u(k))$ of system (12) can be represented by using the constant RBF network $\bar{W}^{s\mathrm{T}}S(x(k))$ along the trajectory φ^s (defined in Section 3). This representation, based on the fundamental information extracted from the trajectory φ^s, is independent of time [22]. It was also showed that the constant RBF networks $\bar{W}^{s\mathrm{T}}S(x(k))$ trained via deterministic learning naturally have a certain ability of generalization [22].

Remark 4 In the literature of NN-based fault diagnosis, neural networks are used to provide powerful modeling tools. Nevertheless, for diagnosis of general nonlinear systems, it is very restrictive to verify a priori that the PE condition is satisfied [14]. Consequently, it is difficult to guarantee that the employed neural networks can truly approximate the system dynamics and fault functions. In this paper, it was shown that both system dynamics for normal and fault modes can be locally accurately approximated through deterministic learning. The sensitivity to small nonlinear faults is increased substantially, since the modeling uncertainties are reduced, and locally-accurate information of nonlinear faults is provided by using the constant NNs.

Note that the NN approximation of $\psi^s(x(k), u(k))$ is accurate only in a local region along the trajectory φ^s. For the region far away from trajectory φ^s, the general fault function $\psi^s(x(k), u(k))$ is not learned. This reveals that both the normal and fault modes of nonlinear discrete-time systems can be represented by using the locally-accurate approximation of their underlying system dynamics along the state trajectories. The local region Ω_{φ^s} is described by [29]

$$\begin{aligned}\Omega_{\varphi^s} :=& \{(x(k), u(k)) | \mathrm{dist}((x(k), u(k)), \varphi^s) < d \\ &\Rightarrow |\bar{W}_i^{s\mathrm{T}}S_i(x(k)) - \psi_i^s(x(k), u(k))| < \xi_i^*\}, \\ & i = 1, \dots, n,\end{aligned} \quad (15)$$

where $\xi_i^* = \mathrm{O}(\epsilon_1) = \mathrm{O}(\epsilon_2)$ is the approximation error within Ω_{φ^s}, d is a positive constant. The representation can be used in a way that whenever the NN input enters the region Ω_{φ^s} again, the RBF network $\bar{W}^{s\mathrm{T}}S(x(k))$ will provide accurate approximation to the dynamics $\psi^s(x(k), u(k))$.

Therefore, by learning the normal system and various fault systems, a bank of the trained system is represented by the corresponding constant RBF networks.

4.2 Fault detection

In this section, residuals generation and the fault decision scheme is presented, then a rigorous detectability analysis of the proposed detection scheme is also provided.

To be specific, the monitored system is described by

$$\begin{aligned}x(k+1) =& f(x(k), u(k)) + \eta(x(k), u(k)) \\ &+\beta(k - k_0)\phi'(x(k), u(k)),\end{aligned} \quad (16)$$

where $\phi'(x(k), u(k))$ represents the deviation of system dynamics due to an unknown fault.

For the monitored system (16), the general fault function is described by

$$\psi'(x(k), u(k)) = \eta(x(k), u(k)) + \beta(k - k_0)\phi'(x(k), u(k)). \quad (17)$$

By utilizing the learned knowledge about various trained systems, a dynamical model is constructed as follows:

$$\begin{aligned}\bar{x}^h(k+1) =& f(x(k), u(k)) + B(\bar{x}^h(k) - x(k)) \\ &+\bar{W}^{h\mathrm{T}}S(x(k)),\end{aligned} \quad (18)$$

where $h = 0, \dots, M$, $\bar{x}^h(k) = [\bar{x}_1^h(k) \ \cdots \ \bar{x}_n^h(k)]^{\mathrm{T}} \in \mathbb{R}^n$ is the state of the dynamical model, $x(k) = [x_1(k) \ \cdots \ x_n(k)]^{\mathrm{T}} \in \mathbb{R}^n$ is the state of monitored system (16), $B = \mathrm{diag}\{b_1, \dots, b_n\}$ is a diagonal matrix which is kept the same for all normal and fault models, with $0 < b_i < 1$. $\bar{W}^{h\mathrm{T}}S(x(k)$ is the learned knowledge of one trained system.

By combining the monitored system (16) and the dynamical model (18), the following residual system is obtained:

$$\tilde{x}_i^h(k+1) = b_i\tilde{x}_i^h(k) + (\bar{W}_i^{hT}S_i(x(k)) \\ -\psi_i'(x(k),u(k))),\ \ i=1,\dots,n, \tag{19}$$

where $\tilde{x}_i^h(k) \triangleq \bar{x}_i^h(k) - x_i(k)$ is the state estimation error (residual), and $\bar{W}_i^{hT}S_i(x(k)) - \psi_i'(x(k),u(k))$ is the difference of system dynamics between the monitored system and the hth estimator.

The following average L_1 norm of residuals is provided for making decision:

$$\|\tilde{x}_i^h(k)\|_1 = \frac{1}{K}\sum_{j=k-K}^{k-1}|\tilde{x}_i^h(j)|,\ \ k \geqslant K > 1, \tag{20}$$

where $K \in Z_+$ is the preset period constant of the monitored system.

Based on the above mentioned average L_1 norm of residual, we have the following fault detection scheme for a class of nonlinear discrete-time systems.

Fault detection decision scheme Compare $\|\tilde{x}_i^s\|_1$ with $\|\tilde{x}_i^0\|_1$, $i=1,\dots,n$. If, for $s \in \{1,\dots,M\}$, there exists some finite time k^s and some $i \in \{1,\dots,n\}$ such that $\|\tilde{x}_i^s(k^s)\|_1 < \|\tilde{x}_i^0(k^s)\|_1$, then the occurrence of a fault is deduced.

The absolute fault detection time instant $K_{\rm d}$ is defined as the first time instant such that $\|\tilde{x}_i^s(K_{\rm d})\|_1 < \|\tilde{x}_i^0(K_{\rm d})\|_1$, for some $K_{\rm d} \geqslant k_0$ and some $i \in \{1,\dots,n\}$, that is, $K_{\rm d} \triangleq \min\{k : \|\tilde{x}_i^s(k)\|_1 < \|\tilde{x}_i^0(k)\|_1\}$ [14]. The fault detection time $k_{\rm d}$ is defined as the difference between the fault occurrence time k_0 and the absolute fault detection time $K_{\rm d}$, i.e., $k_{\rm d} = K_{\rm d} - k_0$ [13,23].

Lemma 2 Consider the monitored system (16), the fault estimator (18) and the residual system (19). For all $s \in \{1,\dots,M\}$, $i \in \{1,\dots,n\}$ and $k \in [K,k_0)$, when no fault occurs, if the following conditions hold:

1) there exists at least one interval $I = [k_{\rm a}, k_{\rm b}-1] \subseteq [k-K, k-1]$ such that

$$\min_{x(k_\tau),u(k_\tau)\in\Omega_{\varphi^s}} \left|\phi_i^s(x(k_\tau),u(k_\tau))\right| \geqslant 2\mu_i,\ \ \forall k_\tau \in I, \tag{21}$$

where $k_{\rm b} > k_{\rm a} > 1$, $\mu_i > \xi_i^*$;

2) b_i satisfies

$$0 < b_i < \Big(\frac{\xi_i^*}{2\mu_i - \xi_i^*}\Big)^{\frac{l}{l-\underline{l}-1}}, \tag{22}$$

where $l = k_{\rm b} - k_{\rm a} > \underline{l} := \dfrac{2\xi_i^*}{\mu_i - \xi_i^*}K$.

Then $\|\tilde{x}_i^0(k)\|_1 < \|\tilde{x}_i^s(k)\|_1$ holds for all $k \in [K,k_0)$.

Proof

Step 1 We prove that $\|\tilde{x}_i^0(k)\|_1 < \dfrac{\xi_i^*}{1-b_i}$ for all $k \in [K,k_0)$.

For the monitored system (16), prior to the occurrence of faults, the general fault function is described by

$$\psi_i'(x(k),u(k)) = \psi_i^0(x(k),u(k)) = \eta_i(x(k),u(k)). \tag{23}$$

In the time period $K \leqslant k < k_0$, the error dynamics with respect to the nominal model satisfies

$$\tilde{x}_i^0(k+1) = b_i\tilde{x}_i^0(k) + [\bar{W}_i^{0T}S_i(x(k)) - \eta_i(x(k),u(k)]. \tag{24}$$

The solution to the above equation is

$$\tilde{x}_i^0(k) = b_i^k\tilde{x}_i^0(0) + \sum_{j=0}^{k-1} b_i^{k-1-j}[\bar{W}_i^{0T}S_i(x(j)) - \eta_i(x(j),u(j)]. \tag{25}$$

From (15), we have

$$|\eta_i(x(k),u(k) - \bar{W}_i^{0T}S_i((x(k),u(k))| < \xi_i^*, \tag{26}$$

where ξ_i^* is given by (15).

By combining (26) and $0 < b_i < 1$, we have

$$|\tilde{x}_i^0(k)| < b_i^k|\tilde{x}_i^0(0)| + \xi_i^*\sum_{j=0}^{k-1} b_i^{k-1-j} \\ < b_i^k|\tilde{x}_i^0(0)| + \frac{\xi_i^*}{1-b_i}. \tag{27}$$

Since $x_i(0)$ is available, we let $\bar{x}_i^0(0) = 0$, $\tilde{x}_i^0(0) = 0$, then, we have $|\tilde{x}_i^0(k)| < \dfrac{\xi_i^*}{1-b_i}$.

Thereby, we have for all $k \geqslant K > 1$,

$$\|\tilde{x}_i^0(k)\|_1 < \frac{1}{K}\sum_{j=k-K}^{k-1}\frac{\xi_i^*}{1-b_i} = \frac{\xi_i^*}{1-b_i}. \tag{28}$$

Step 2 We prove that $\|\tilde{x}_i^s(k)\|_1 > \dfrac{\xi_i^*}{1-b_i} > \|\tilde{x}_i^0(k)\|_1$ for all $k \in [K,k_0)$.

The error dynamics with respect to the sth model satisfies

$$\tilde{x}_i^s(k+1) = b_i\tilde{x}_i^s(k) + [\bar{W}_i^{sT}S_i(x(k)) - \eta_i(x(k),u(k)]. \tag{29}$$

Thus, for $k_\tau \in I$,

$$|\tilde{x}_i^s(k_\tau)| = |b_i^{k_\tau-k_{\rm a}}\tilde{x}_i^s(k_{\rm a}) \\ +\sum_{j=k_{\rm a}}^{k_\tau-1} b_i^{k_\tau-1-j}[\bar{W}_i^{sT}S_i(x(j))-\eta_i(x(j),u(j))]|. \tag{30}$$

From (15) and (21), we have

$$\begin{aligned}
&|\bar{W}_i^{s\mathrm{T}} S_i(x(k_\tau)) - \eta_i(x(k_\tau), u(k_\tau)| \\
&\geqslant |\phi_i^s(x(k_\tau), u(k_\tau))| - |\bar{W}_i^{s\mathrm{T}} S_i(x(k_\tau)) \\
&\quad -\phi_i^s(x(k_\tau), u(k_\tau)) - \eta_i(x(k_\tau), u(k_\tau)| \\
&\geqslant 2\mu_i - \xi_i^* \\
&> \mu_i.
\end{aligned} \tag{31}$$

Let $I' = [k'_\mathrm{a}, k'_\mathrm{b} - 1] \subseteq I$ with $k'_\mathrm{b} - 1 > k'_\mathrm{a}$ and I' be defined as

$$I' = \{k_\tau \in I : |\tilde{x}_i^s(k_\tau)| < \frac{\mu_i - \xi_i^*}{1 - b_i}\}. \tag{32}$$

Let l, l' denote the length of the time interval I and I' respectively, thereby $l' = k'_\mathrm{b} - k'_\mathrm{a} > 1$, $l \geqslant l'$.

Then, it can be proven that there at most exists one time interval I' and $l' \leqslant \dfrac{\ln \frac{b_i\xi_i^*}{2\mu_i - \xi_i^*}}{\ln b_i}$.

The magnitude of $|\tilde{x}_i^s(k_\tau)|$ in the time interval I can be discussed in the following cases:

1) If $|\tilde{x}_i^s(k_\mathrm{a})| \geqslant \dfrac{\mu_i - \xi_i^*}{1 - b_i}$, and $[\bar{W}_i^{s\mathrm{T}} S_i(x(k)) - \eta_i(x(k), u(k))]$ has the same sign with $\tilde{x}_i^s(k_\mathrm{a})$, then from (30) and (31), we have

$$\begin{aligned}
|\tilde{x}_i^s(k_\tau)| &= |b_i^{k_\tau - k_\mathrm{a}} \tilde{x}_i^s(k_\mathrm{a}) \\
&\quad + \sum_{j=k_\mathrm{a}}^{k_\tau - 1} b_i^{k_\tau - 1 - j} [\bar{W}_i^{s\mathrm{T}} S_i(x(j)) - \eta_i(x(j), u(j))]| \\
&\geqslant |b_i^{k_\tau - k_\mathrm{a}} \tilde{x}_i^s(k_\mathrm{a}) + \mu_i \frac{1 - b_i^{k_\tau - k_\mathrm{a}}}{1 - b_i}| \\
&\geqslant \frac{\mu_i - (\xi_i^* + \epsilon_i^*)}{1 - b_i} b_i^{k_\tau - k_\mathrm{a}} + \frac{\mu_i}{1 - b_i} - \frac{\mu_i}{1 - b_i} b_i^{k_\tau - k_\mathrm{a}} \\
&\geqslant \frac{\mu_i}{1 - b_i} - \frac{\xi_i^*}{1 - b_i} \\
&= \frac{\mu_i - \xi_i^*}{1 - b_i}.
\end{aligned} \tag{33}$$

Thus, $|\tilde{x}_i^s(k_\tau)| \geqslant \dfrac{\mu_i - \xi_i^*}{1 - b_i}$ holds for all $k_\tau \in I$. Therefore, $I' = \emptyset$ and $l' = 0$.

2) If $|\tilde{x}_i^s(k_\mathrm{a})| < \dfrac{\mu_i - \xi_i^*}{1 - b_i}$, according to (30) and (31), we have for all $k_\tau \in I$

$$\begin{aligned}
|\tilde{x}_i^s(k_\tau)| &= |b_i^{k_\tau - k_\mathrm{a}} \tilde{x}_i^s(k_\mathrm{a}) \\
&\quad + \sum_{j=k_\mathrm{a}}^{k_\tau - 1} b_i^{k_\tau - 1 - j} [\bar{W}_i^{s\mathrm{T}} S_i(x(j)) - \eta_i(x(j), u(j))]| \\
&\geqslant |\mu_i \frac{1 - b_i^{k_\tau - k_\mathrm{a}}}{1 - b_i}| - |b_i^{k_\tau - k_\mathrm{a}} \tilde{x}_i^s(k_\mathrm{a})| \\
&\geqslant |\mu_i \frac{1 - b_i^{k_\tau - k_\mathrm{a}}}{1 - b_i}| - |b_i^{k_\tau - k_\mathrm{a}} \frac{\mu_i - \xi_i^*}{1 - b_i}| \\
&= \frac{\mu_i}{1 - b_i} - \frac{\mu_i}{1 - b_i} b_i^{k_\tau - k_\mathrm{a}} - \frac{\mu_i - \xi_i^*}{1 - b_i} b_i^{k_\tau - k_\mathrm{a}} \\
&= \frac{\mu_i}{1 - b_i} - \frac{2\mu_i - \xi_i^*}{1 - b_i} b_i^{k_\tau - k_\mathrm{a}}.
\end{aligned} \tag{34}$$

Consider (34), if $k'_\mathrm{a} = k_\mathrm{a}$, then there exists a finite time instant $k'_\mathrm{b} - 1$ such that

$$\frac{\mu_i - \xi_i^*}{1 - b_i} \geqslant |\tilde{x}_i^s(k'_\mathrm{b} - 1)| \geqslant \frac{\mu_i}{1 - b_i} - \frac{2\mu_i - \xi_i^*}{1 - b_i} b_i^{k'_\mathrm{b} - 1 - k'_\mathrm{a}}. \tag{35}$$

Solving the above inequality for k'_b yields

$$k'_\mathrm{b} \leqslant \frac{\ln \frac{\xi_i^*}{2\mu_i - \xi_i^*}}{\ln b_i} + 1 + k'_\mathrm{a}, \tag{36}$$

which implies that

$$l' = k'_\mathrm{b} - k'_\mathrm{a} \leqslant \frac{\ln \frac{\xi_i^*}{2\mu_i - \xi_i^*}}{\ln b_i} + 1 = \frac{\ln \frac{b_i\xi_i^*}{2\mu_i - \xi_i^*}}{\ln b_i}. \tag{37}$$

Thus we have $|\tilde{x}_i^s(k_\mathrm{b} - 1)| \geqslant \dfrac{\mu_i - \xi_i^*}{1 - b_i}$ and $|\tilde{x}_i^s(k'_\mathrm{b} - 1)| \geqslant \dfrac{\mu_i - \xi_i^*}{1 - b_i}$, it can be proved that $[\bar{W}_i^{s\mathrm{T}} S_i(x(k)) - \eta_i(x(k), u(k))]$ has the same sign with $\tilde{x}_i^s(k'_\mathrm{b} - 1)$. By using the analysis result of case 1), we have $|\tilde{x}_i^s(k_\tau)| \geqslant \dfrac{\mu_i - \xi_i^*}{1 - b_i}$ holds for all $k_\tau \in [k'_\mathrm{b}, k_\mathrm{b} - 1]$. Therefore, $I' = [k'_\mathrm{a}, k'_\mathrm{b} - 1]$ and $l' \leqslant \dfrac{\ln \frac{b_i\xi_i^*}{2\mu_i - \xi_i^*}}{\ln b_i}$.

3) In the case that $|\tilde{x}_i^s(k_\mathrm{a})| \geqslant \dfrac{\mu_i - \xi_i^*}{1 - b_i}$, and $[\bar{W}_i^{s\mathrm{T}} S_i(x(k)) - \eta_i(x(k), u(k))]$ has a different sign with $\tilde{x}_i^s(k_\mathrm{a})$, if there exists a time $k'_\mathrm{a} \in I$ such that $|\tilde{x}_i^s(k'_\mathrm{a})| < \dfrac{\mu_i - \xi_i^*}{1 - b_i}$, then using analysis result of case 2), we have $I' = [k'_\mathrm{a}, k'_\mathrm{b} - 1]$ and $l' \leqslant \dfrac{\ln \frac{b_i\xi_i^*}{2\mu_i - \xi_i^*}}{\ln b_i}$. If such time k'_a does not exist, then $I' = \emptyset$ and $l' = 0$.

From the above discussion, we can summarize that there at most exists one time interval $I' = [k'_\mathrm{a}, k'_\mathrm{b} - 1] \subseteq I$ and $l' \leqslant \dfrac{\ln \frac{b_i\xi_i^*}{2\mu_i - \xi_i^*}}{\ln b_i}$.

Thus, in light of (20), we have for all $k \geqslant K > 1$,

$$\begin{aligned}
&\|\tilde{x}_i^s(k)\|_1 \\
&= \frac{1}{K}\sum_{j=k-K}^{k-1} |\tilde{x}_i^s(j)| \\
&\geqslant \frac{1}{K}[\sum_{j=k_a}^{k'_a-1} |\tilde{x}_i^s(j)| + \sum_{j=k'_a}^{k'_b-1} |\tilde{x}_i^s(j)| + \sum_{j=k'_b}^{k_b-1} |\tilde{x}_i^s(j)|] \\
&\geqslant \frac{1}{K}[\sum_{j=k_a}^{k'_a-1} |\tilde{x}_i^s(j)| + \sum_{j=k'_b}^{k_b-1} |\tilde{x}_i^s(j)|] \\
&\geqslant \frac{1}{K}[(k_b - k_a) - (k'_b - k'_a)]\frac{\mu_i - (\xi_i^* + \epsilon_i^*)}{1 - b_i} \\
&= \frac{1}{K}(l - l')\frac{\mu_i - \xi_i^*}{1 - b_i} \\
&\geqslant \frac{\underline{l}(\mu_i - \xi_i^*)}{K(1 - b_i)}. \qquad (38)
\end{aligned}$$

With (28) and $0 < b_i < 1$, (38) can be written as

$$\|\tilde{x}_i^s(k)\|_1 \geqslant \frac{2\xi_i^*}{1 - b_i} > \frac{\xi_i^*}{1 - b_i} > \|\tilde{x}_i^0(k)\|_1. \qquad (39)$$

□

The following theorem characterizes the fault detectability properties which include the fault detectability condition and detection time.

Theorem 1 Consider the monitored system (16), the fault estimator (18) and the residual system (19). For some $s \in \{1, \dots, M\}$, some $i \in \{1, \dots, n\}$ and $k \geqslant k_0 + K$, if the following conditions hold:

1) there exists at least one interval $I = [k_a, k_b - 1] \subseteq [k - K, k - 1]$ such that

$$|\phi'_i(x(k_\tau), u(k_\tau))| \geqslant 2\mu_i, \quad \forall k_\tau \in I, \qquad (40)$$

where $k_b > k_a > 1$, $\mu_i > \xi_i^*$.

2) b_i satisfies

$$0 < b_i < (\frac{\xi_i^*}{2\mu_i - \xi_i^*})^{\frac{l}{l - \underline{l} - 1}}, \qquad (41)$$

where $l = k_b - k_a > \underline{l} := \dfrac{2\xi_i^*}{\mu_i - \xi_i^*}K$.

Then,

1) the fault will be detected. i.e., there exists a finite time k^s such that $\|\tilde{x}_i^s(k^s)\|_1 < \|\tilde{x}_i^0(k^s)\|_1$;

2) the upper bound on the detection time k_d is given by

$$\bar{k}_d = \min_{i \in \{1,\dots,n\}} \log_{b_i} \frac{\xi_i^* K}{(1 - b_i^K)|\tilde{x}_i^s(k_0)|} + K. \qquad (42)$$

Proof The error dynamics with respect to the nominal model, after the occurrence of the fault yield the following form:

$$\tilde{x}_i^0(k+1) = b_i\tilde{x}_i^0(k) + [\bar{W}_i^{0\mathrm{T}} S_i(x(k)) - \psi'_i(x(k), u(k)]. \qquad (43)$$

For any $k \geqslant k_0 + K$, the solution to (48) is given by

$$\begin{aligned}
\tilde{x}_i^0(k) =& b_i^{k-k_0}\tilde{x}_i^0(k_0) \\
&+ \sum_{j=k_0}^{k-1} b_i^{k-1-j}[\bar{W}_i^{0\mathrm{T}} S_i(x(j)) - \psi'_i(x(j), u(j))]. \qquad (44)
\end{aligned}$$

For $k_\tau \in I$, using the triangle inequality and (40), we have

$$\begin{aligned}
&|\bar{W}_i^{0\mathrm{T}} S_i(x(k_\tau)) - \psi'_i(x(k), u(k))| \\
&= |\phi'_i(x(k_\tau), u(k_\tau))| - |\bar{W}_i^{0\mathrm{T}} S_i(x(k_\tau)) - \eta_i(x(k_\tau), u(k_\tau))| \\
&\geqslant 2\mu_i - \xi_i^* \\
&> \mu_i. \qquad (45)
\end{aligned}$$

Following the same steps in the proof of Lemma 2, let I' be defined as

$$I' = \{k_\tau \in I : |\tilde{x}_i^0(k_\tau)| < \frac{\mu_i - \xi_i^*}{1 - b_i}\} \qquad (46)$$

and l, l' denote the length of the time interval I and I' respectively.

When $k \geqslant k_0 + T$, from (38), we have

$$\begin{aligned}
\|\tilde{x}_i^0(k)\|_1 &= \frac{1}{K}\sum_{j=k-K}^{k-1} |\tilde{x}_i^0(j)| \\
&\geqslant \frac{1}{K}\sum_{j=k_a}^{k_b-1} |\tilde{x}_i^0(j)| \geqslant \frac{2\xi_i^*}{1 - b_i}. \qquad (47)
\end{aligned}$$

The error dynamics with respect to the sth model satisfies

$$\tilde{x}_i^s(k+1) = b_i\tilde{x}_i^s(k) + [\bar{W}_i^{s\mathrm{T}} S_i(x(k)) - \psi'_i(x(k), u(k))]. \qquad (48)$$

Thus, when the fault occurs, the solution to the above equation is

$$\begin{aligned}
\tilde{x}_i^s(k) =& b_i^{k-k_0}\tilde{x}_i^s(k_0) \\
&+ \sum_{j=k_0}^{k-1} b_i^{k-1-j}[\bar{W}_i^{s\mathrm{T}} S_i(x(j)) - \psi'_i(x(j), u(j))]. \qquad (49)
\end{aligned}$$

From(15), we have

$$\begin{aligned}
|\tilde{x}_i^s(k)| \leqslant& |b_i^{k-k_0}\tilde{x}_i^s(k_0) \\
&+ \sum_{j=k_0}^{k-1} b_i^{k-1-j}[\bar{W}_i^{s\mathrm{T}} S_i(x(j)) - \psi'_i(x(j), u(j))]|
\end{aligned}$$

$$< \xi_i^* \sum_{j=k_0}^{k-1} b_i^{k-1-j} + b_i^{k-k_0}|\tilde{x}_i^s(k_0)|$$
$$< \frac{\xi_i^*}{1-b_i} + b_i^{k-k_0}|\tilde{x}_i^s(k_0)|. \tag{50}$$

Thus, for $k \geqslant k_0 + K$

$$\begin{aligned}\|\tilde{x}_i^s(k)\|_1 &= \frac{1}{K}\sum_{j=k-K}^{k-1}|\tilde{x}_i^s(j)| \\ &< \frac{1}{K}\sum_{j=k-K}^{k-1}\frac{\xi_i^*}{1-b_i} + \frac{1}{K}\sum_{j=k-K}^{k-1} b_i^{j-k_0}|\tilde{x}_i^s(k_0)| \\ &= \frac{\xi_i^*}{1-b_i} + \frac{1}{K}|\tilde{x}_i^s(k_0)|\sum_{j=k-K}^{k-1} b_i^{j-k_0} \\ &= \frac{\xi_i^*}{1-b_i} + \frac{b_i^{k-K-k_0}(1-b_i^K)}{(1-b_i)K}|\tilde{x}_i^s(k_0)|.\end{aligned} \tag{51}$$

From (47) and (51), we have $\|\tilde{x}_i^s(k)\|_1 < \|\tilde{x}_i^0(k)\|_1$ if

$$\frac{\xi_i^*}{1-b_i} + \frac{|\tilde{x}_i^s(k_0)|}{(1-b_i)K}(1-b_i^K)b_i^{k-K-k_0} < \frac{2\xi_i^*}{1-b_i}. \tag{52}$$

Note that the left-hand side of the aforementioned inequality is a decreasing function of k. This implies there exists a finite time k^s such that the aforementioned inequality is satisfied [23].

2) The fault detection time k_{d} can be obtained by solving the following equation:

$$\frac{|\tilde{x}_i^s(k_0)|}{(1-b_i)K}(1-b_i^K)b_i^{k-K-k_0} < \frac{\xi_i^*}{1-b_i}. \tag{53}$$

Then, we obtain

$$k > \log_{b_i}\frac{\xi_i^* + K}{(1-b_i^K)|\tilde{x}_i^s(k_0)|} + k_0 + K. \tag{54}$$

Since $k \geqslant k_0 + K$, the absolute fault detection time instant K_{d} satisfies

$$\begin{aligned}K_{\mathrm{d}} &< \max\{k_0 + K, \min_{i=1,\dots,n}\log_{b_i}\frac{\xi_i^* K}{(1-b_i^K)|\tilde{x}_i^s(k_0)|} + k_0 + K\} \\ &< \min_{i=1,\dots,n}\log_{b_i}\frac{\xi_i^* K}{(1-b_i^K)|\tilde{x}_i^s(k_0)|} + k_0 + K.\end{aligned} \tag{55}$$

The proof is completed by letting $k_{\mathrm{d}} = K_{\mathrm{d}} - k_0$. □

Remark 5 As shown in Lemma 2 and Theorem 1, increasing b_i in equation (18) will improve the detectability properties of the presented approach and decrease the detection time. On the other hand, increasing b_i results in increased stiffness of the differential equation, which may lead to numerical problems and increase computational time [13]. Therefore, b_i should be chosen sufficiently large to ensure the adequate rapidness and sensitivity of the detection scheme.

Remark 6 K is the preset period constant of the monitored system in equations (20)–(22). In general, the value of K can be set as the multiples of the monitored system period. The interval $\{k_{\mathrm{a}}, \dots, k_{\mathrm{b}}\}$ represents a time segment after the transient process in modeling phase. The length of interval $I = [k_{\mathrm{a}}, k_{\mathrm{b}}-1] \subseteq [k-K, k-1]$ must be satisfied.

5 Simulation

To show the effectiveness and efficiency of the proposed scheme, the well-known three-tank problem is considered [41]. It contains three interconnected water tanks, two pumps and associated valves as shown in Fig. 1.

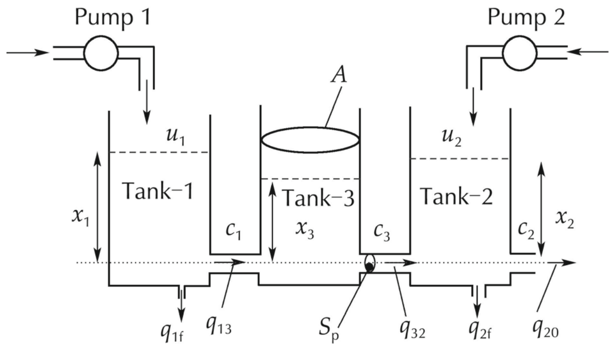


Fig. 1 Structure of the three-tank system under consideration.

The discrete-time model will be obtained from the literature [14] by employing a simple forward Euler discretization, with $T_s = 0.001$:

$$\begin{cases} x_1(k+1) = x_1(k) + \dfrac{T_s}{A}(-q_{13}(k) + u_1(k)), \\ x_2(k+1) = x_2(k) + \dfrac{T_s}{A}(q_{32}(k) - q_{20}(k) + u_2(k)), \\ x_3(k+1) = x_3(k) + \dfrac{T_s}{A}(q_{13}(k) - q_{32}(k)), \end{cases} \tag{56}$$

where k is the discrete-time instant.

The variables used in this system are given below:

$x_i(k)$: Tank i liquid level (m), and $0 < x_i(k) < 0.69$ m, $i = 1, 2, 3$.

A: The cross section (m^2) of tanks, and all the tanks have $A = 0.0154$ m^2.

$q_{13}(k)$: The fluid flow rate (m^3/s) between tanks 1 and 3, $q_{13}(k) = c_1 S_{\mathrm{p}}\mathrm{sgn}(x_1(k) - x_3(k))\sqrt{2\mathrm{g}|x_1(k) - x_3(k)|}$.

$q_{32}(k)$: The fluid flow rate (m^3/s) between tanks 3 and 2, $q_{32}(k) = c_3 S_p \mathrm{sgn}(x_3(k) - x_2(k)) \sqrt{2g|x_3(k) - x_2(k)|}$.

$q_{20}(k)$: The fluid flow of outlet rate (m^3/s) from the tank 2, $q_{20}(k) = c_2 S_p \sqrt{2gx_2(k)}$.

c_i: The outflow coefficients, $c_1 = 1$, $c_2 = 0.8$, $c_3 = 1$.

S_p: The cross section (m^2) of the connection pipes, $S_p = 5 \times 10^{-5}\,m^2$.

g: The gravity acceleration (m/s^2), $g = 9.8\,m/s^2$.

$u_i(k)$: The fluid flow of inlet rate (m^3/s) from two pumps, and $0 \leqslant u_i(k) \leqslant 1.2 \times 10^{-4}\,m^3/s$, $i = 1, 2$. Moreover, $u_i(k) = \Psi(v_i(k))$, where $v_1(k) = -5S_p(x_1(k) - 0.5) + 0.8S_p(1.5 + \sin(\omega_1 k))$, $v_2(k) = -5S_p(x_2(k) - 0.5) + 0.8S_p(1.5 + \cos(\omega_2 k))$, $\omega_1 = 0.3$, $\omega_2 = 0.3$. $\Psi(v_i(k)) = 0$ if $v_i < 0$ or $v_i > 1.2 \times 10^{-4}\,m^3/s$, $\Psi(v_i(k)) = v_i(k)$ if $0 \leqslant v_i \leqslant 1.2 \times 10^{-4}\,m^3/s$.

The phase portrait of the system (56) in normal mode is shown in Fig. 2. The state responses of the normal mode is seen in Fig. 3.

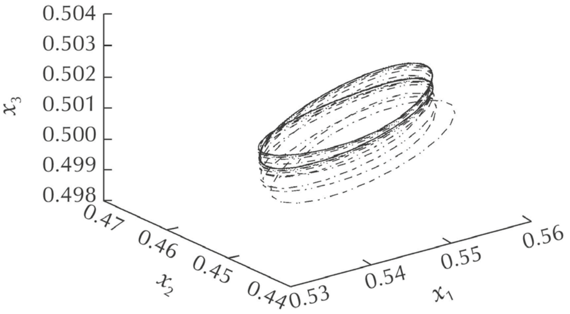


Fig. 2 The phase trajectory of the three-tank system in space.

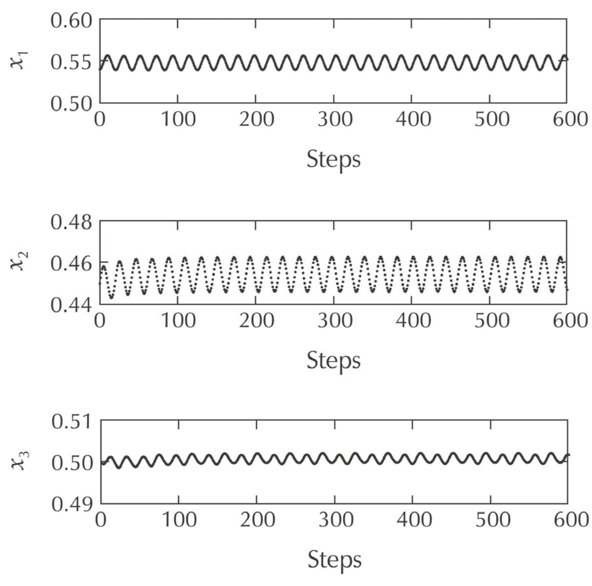


Fig. 3 Water levels $x_i(k)$.

5.1 Fault modes of three-tanks system

In the three-tank system simulation, the control input terms $u_i(k)$ are the fluid flow of inlet rate from two pumps. Water levels $x_i(k)$ are in oscillations. The amplitudes of the oscillations are small, which are only about 0.01 m. This situation is common in practical application. The purpose of the simulation is to show that accurate NN approximations of the nonlinear system dynamics can be achieved in the local regions near the system equilibrium points. The local NN approximations are then used to detect the faults.

We consider four types of abrupt faults, and also hypothesize every type contains only one fault. The trajectory of fault is close to the normal trajectory when the fault occurs. Moreover, it is assumed that the measurements of the water levels $x_i(k)$, which are influenced by a pump fault or a leakage. The behaviour of the system (56) in the following situations:

1) Normal mode: The system is not faulty.

In this simulation, the $x_i(k+1)$ subsystem is considered to be unknown, i.e., state equations represent the general fault function of normal mode.

2) Fault mode 1: Actuator fault in pump 1.

We consider a simple multiplicative actuator fault in pump 1 by letting $u_1'(k) = u_1(k) + (\alpha_1 - 1)u_1(k)$, where $u_1(k)$ is the supply flow rate in the non-fault case, and α_1 is the parameter characterizing the magnitude of the fault. For $\alpha_1 = 1$, we have the non-fault situation in pump 1, whereas $\alpha_1 = 0$ implies that the pump is completely faulty, in the sense that there is no flow. For this simulation, $\alpha_1 = 0.8$.

3) Fault mode 2: Actuator fault in pump 2.

Analogously to the case of a fault in pump 1, we have $u_2'(k) = u_2(k) + (\alpha_2 - 1)u_2(k)$, $\alpha_2 = 0.8$.

4) Fault mode 3: Leakage in tank 1.

We assume that the leak is circular in shape and of cross section $S_{l1} = 4 \times 10^{-5}\,m^2$. Then, the outflow rate of the leak is $q_{1f} = c_1 S_{l1} \sqrt{2gx_1(k)}$ in tank 1.

5) Fault mode 4: Leakage in tank 2.

Analogously to the case of a leakage in tank 1, we have $S_{l2} = 8 \times 10^{-5}\,m^2$, $q_{2f} = c_2 S_{l2} \sqrt{2gx_2(k)}$.

These faults are presented only as representative examples of all the countless faults which can take place in the system. The case of incipient faults is completely analogous and is not addressed here for the sake of brevity.

Consider the same initial condition is $x(0) = [x_1(0)\ x_2(0)\ x_3(0)]^T = [0.54\ 0.45\ 0.5]^T$. The phase portrait and state responses of the fault modes are seen in Figs. 4–11.

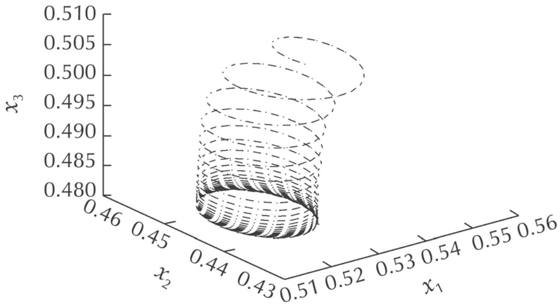


Fig. 4 The phase trajectory of the fault mode 1.

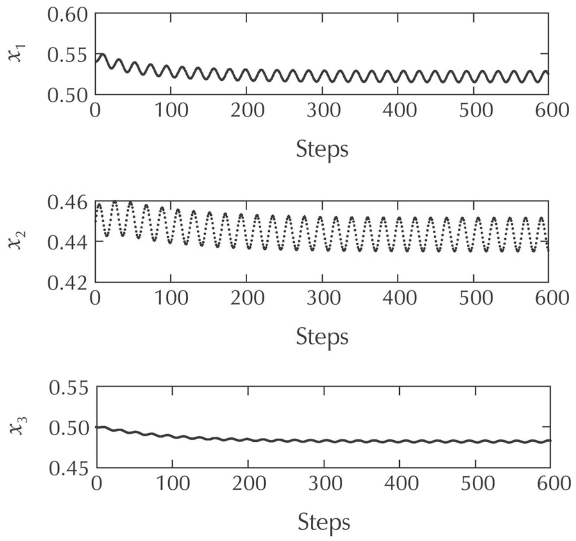


Fig. 5 Water levels $x_i(k)$ of the fault mode 1.

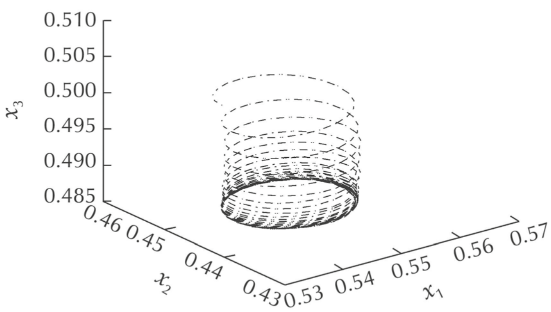
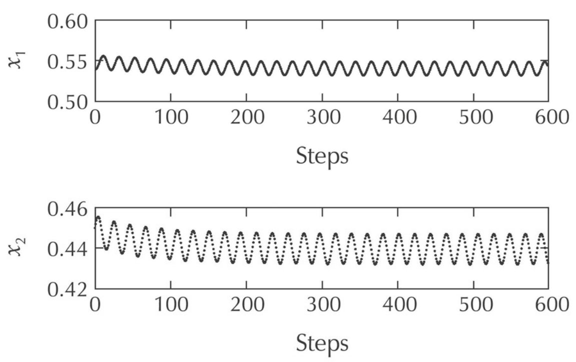


Fig. 6 The phase trajectory of the fault mode 2.



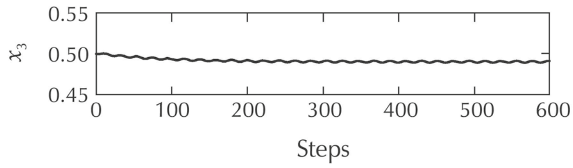


Fig. 7 Water levels $x_i(k)$ of the fault mode 2.

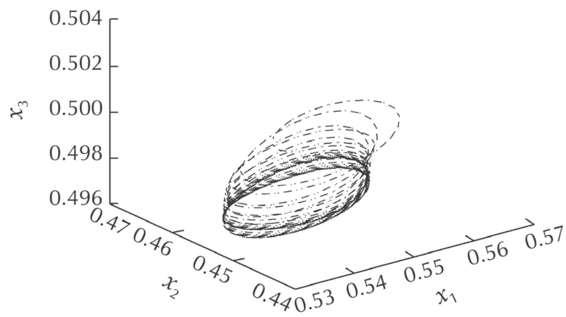


Fig. 8 The phase trajectory of the fault mode 3.

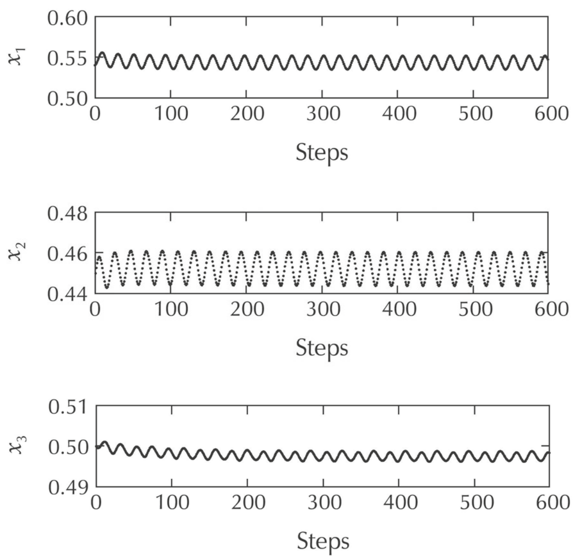


Fig. 9 Water levels $x_i(k)$ of the fault mode 3.

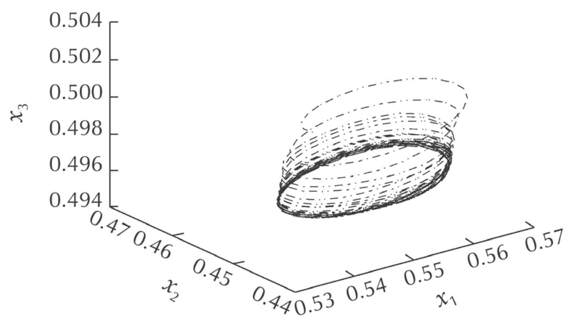


Fig. 10 The phase trajectory of the fault mode 4.

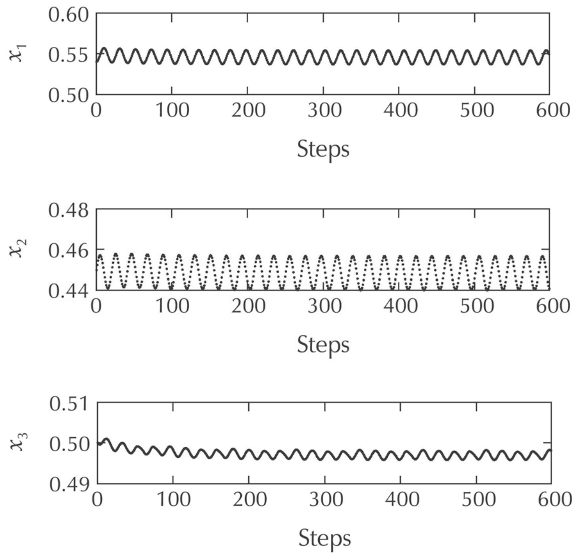


Fig. 11 Water levels $x_i(k)$ of the fault mode 4.

5.2 Training phase

For the discrete-time system (56), the following dynamical RBF networks are employed to learn the normal and fault modes:

$$\begin{cases} \hat{x}_i(k) = a_i(z_i(k-1) - x_i(k-1)) \\ \qquad\quad +\hat{W}_i^{s\mathrm{T}}(k)S_i(x(k-1)), \\ z_i(k) = a_i(z_i(k-1) - x_i(k-1)) \\ \qquad\quad +\hat{W}_i^{s\mathrm{T}}(k+1)S_i(x(k-1)), \\ i = 1, 2, \ \ s = 0, \dots, 4. \end{cases} \tag{57}$$

The RBF network $\hat{W}_1^{s\mathrm{T}} S_1(x(k))$ with nodes $N = 18 \times 11 \times 28$ nodes are constructed to learn the unknown system dynamic in the first state equation in normal and fault modes, whereas the RBF network $\hat{W}_2^{s\mathrm{T}} S_2(x(k))$ with the same nodes are constructed to learn the system dynamic in the second state equation in normal and fault modes, $s = 0, \dots, 4$. The centers of RBF network evenly spaced on $[0.510, 0.580] \times [0.430, 0.465] \times [0.475, 0.505]$ with the widths $\eta = 0.0025$. Figs. 3, 5, 7, 9 and 11 show the NN inputs $x_1(k)$, $x_2(k)$ and $x_3(k)$ for training the normal and fault modes, respectively.

Parameters in equation (57) and the weights updated law (6) of RBF network are chosen as $a_1 = a_2 = 0.5$, $\Gamma_1 = \Gamma_2 = 2$, respectively. The initial condition is set as $\hat{W}_i^s(0) = 0$.

Consider the actuator fault of fault mode 1, with the initial condition is $x(0) = [x_1(0) \ \ x_2(0) \ \ x_3(0)]^{\mathrm{T}} = [0.54 \ \ 0.45 \ \ 0.50]^{\mathrm{T}}$, fix the system parameter as $\alpha_1 = 0.8$ to generate a periodic discrete-time sequence. Using the dynamical RBF networks (57) to learn the general fault functions, Fig. 12 displays the trajectory of fault mode 1 and NN coverage in space.

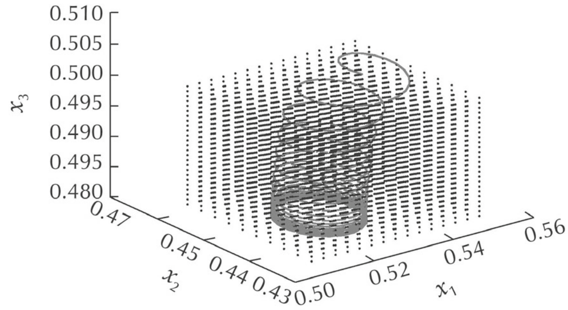


Fig. 12 NN coverage of the fault mode 1.

Fig. 13 show that convergence of partial NN parameter to their optimal values and locally accurate approximation of system dynamics are implemented in the process of training fault modes. The learned knowledge is stored in constant RBF networks $\bar{W}_1^{1\mathrm{T}} S_1(x(k))$.

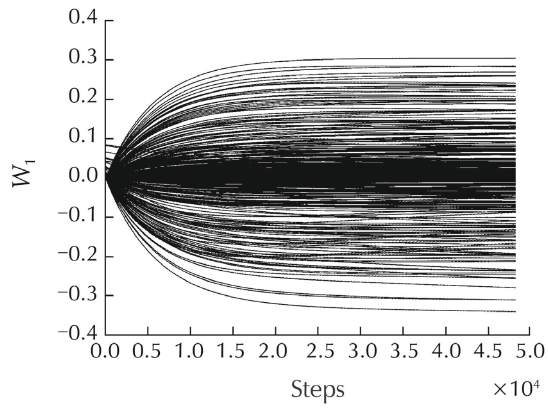


Fig. 13 Partial NN parameter convergence of $\hat{W}_1^1(k)$.

Simulation results for training of other faults and modes are similar to the above fault mode, and are omitted here due to the limitation of space.

6 Fault detection phase

We consider two types of monitored system with abrupt fault, and also hypothesize every type contains only one fault.

Assume that for the monitored system 1, a fault occurs at $k_0 = 250$ steps due to the multiplicative actuator fault in pump 1. the parameter of the fault magnitude is $\alpha_1' = 0.85$. The phase trajectory and states of the monitored system 1 are shown in Figs. 14 and 15.

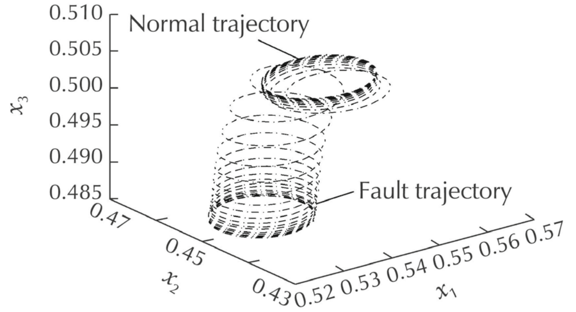


Fig. 14 The phase trajectory of the monitored system 1.

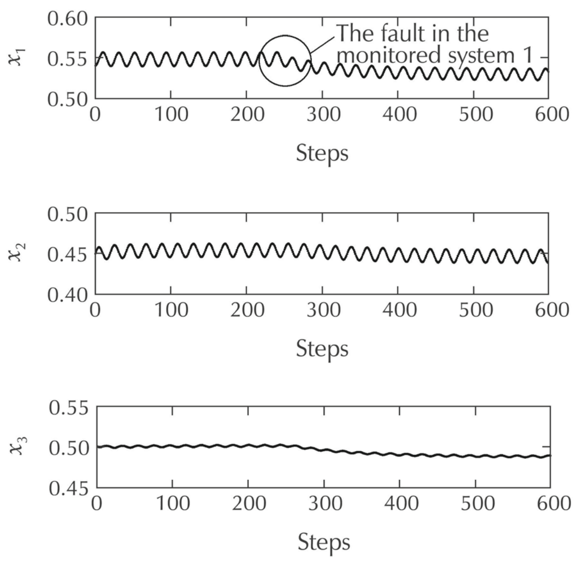


Fig. 15 Water levels $x_i(k)$ of the monitored system 1.

Assume that for the monitored system 2, a leakage fault occurs at $k_0 = 250$ steps with cross section $S'_{l1} = 5\times10^{-6}\,\mathrm{m}^2$ in tank 1. The phase trajectory and states of the monitored system 2 are shown in Figs. 16 and 17. The process for learning other actuator and plant faults are similar to those detecting results of the monitored systems 1 and 2, and are omitted here.

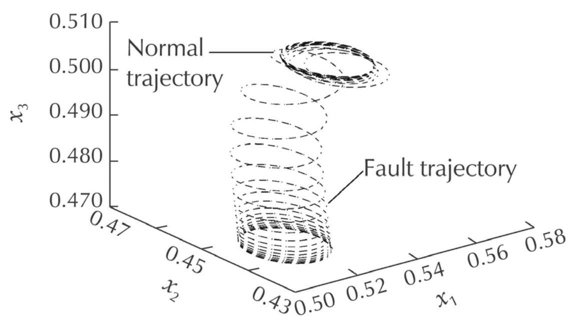


Fig. 16 The phase trajectory of the monitored system 2.

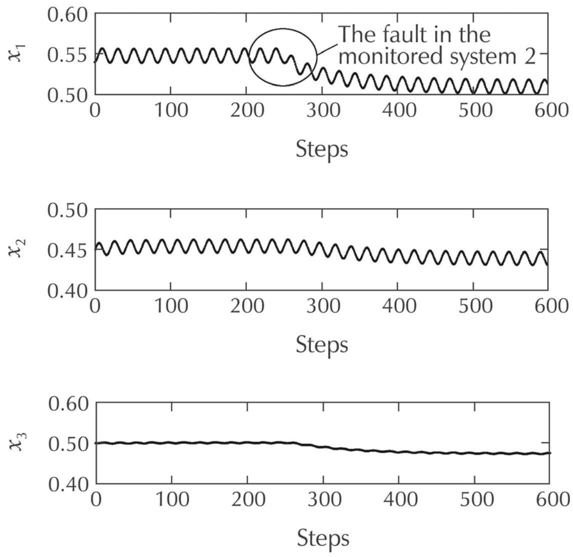


Fig. 17 Water levels $x_i(k)$ of the monitored system 2.

Consider fault modes 0, 1 and 3 because of the monitored fault occurs at pump 1. Three RBF NN estimators are constructed for detecting the first state equation of the monitored system:

$$\bar{x}_1^h(k+1) = b_1(\bar{x}_1^h(k) - x_1(k)) + \bar{W}_1^{h\mathrm{T}} S(x(k)), \quad h = 0,1,3, \tag{58}$$

whereas three RBF NN estimators are constructed for detecting the second state equation of the monitored system:

$$\bar{x}_2^h(k+1) = b_2(\bar{x}_2^h(k) - x_2(k)) + \bar{W}_2^{h\mathrm{T}} S(x(k)), \quad h = 0,1,3, \tag{59}$$

where $\bar{x}_1^h$, $\bar{x}_2^h$ are the state of the estimators, x_1 and x_2 are the state of the monitored systems. The parameter $b_1 = b_2 = 0.5$ is designed. $\tilde{x}_1^h(k) = \bar{x}_1^h(k) - x_1(k)$ and $\tilde{x}_2^h(k) = \bar{x}_2^h(k) - x_2(k)$ are the residuals.

According to (20), the average L_1 norm of the residuals is calculated with $K = 30$ steps. The detection time is yielded because of the choice of the preset period constant of the monitored system (i.e., K) for calculating the average L_1 norm of the residuals.

For the first state equation of the monitored system 1, the average L_1 norm $\|\tilde{x}_1^s\|_1 (s = 0,1,3)$ are shown in Fig. 18. $\|\tilde{x}_1^1\|_1$ becomes smaller than $\|\tilde{x}_1^3\|_1$ at approximately $K_\mathrm{d} = 319$ steps. It is mean that the test fault is detected at $K_\mathrm{d} = 319$ steps, and the fault detection time $k_\mathrm{d} = K_\mathrm{d} - k_0 = 69$ steps.

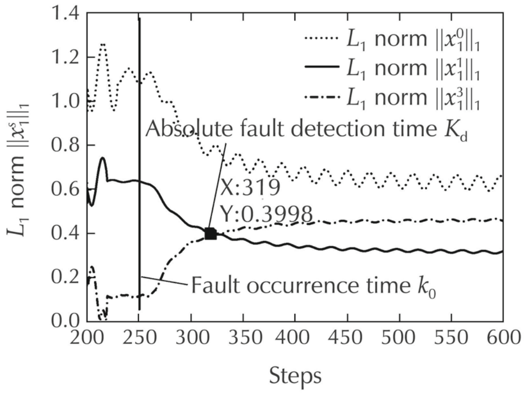


Fig. 18 The average L_1 norm $\|\tilde{x}_1^s\|_1 (s = 0, 1, 3)$ of the monitored system 1.

For the first state equation of the monitored system 2, the average L_1 norm $\|\tilde{x}_1^s\|_1$ $(s = 0, 1, 3)$ are shown in Fig. 19. $\|\tilde{x}_1^1\|_1$ becomes smaller than $\|\tilde{x}_1^3\|_1$ at approximately $K_d = 290$ steps. Thus the test fault is detected at $K_d = 290$ steps, and the fault detection time $k_d = K_d - k_0 = 40$ steps.

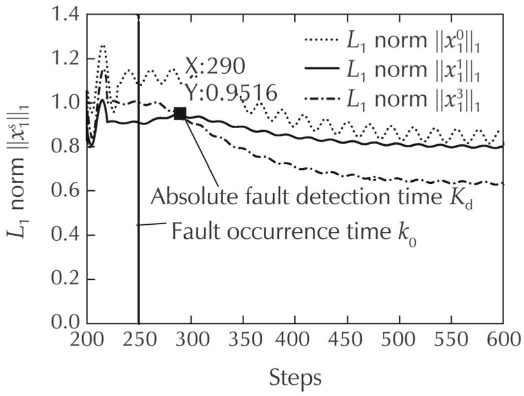


Fig. 19 The average L_1 norm $\|\tilde{x}_1^s\|_1 (s = 0, 1, 3)$ of the monitored system 2.

7 Conclusions

In this paper, an approach has been proposed for fault detection of nonlinear discrete-time systems based on the recently proposed deterministic learning theory. First, the discrete-time dynamical RBF network with an extended weight update law has been employed to locally-accurately approximate the general fault function (i.e., the internal dynamic) via deterministic learning. The obtained knowledge of system dynamics has been embedded to construct a bank of estimators, and a set of residuals have been obtained and used to measure the differences between the dynamics of the monitored system and the dynamics of the trained systems. Second, a fault detection decision scheme has been presented by comparing the magnitude of residuals. Subsequently, the fault detectability analysis has been carried out and the upper bound of detection time has been derived. Finally, a simulation example has been given to illustrate the effectiveness of the proposed scheme.

Future research effort will be devoted to address several issues including: i) performance analysis of the proposed fault detection scheme; ii) fault isolation for nonlinear discrete-time systems; iii) fault detection and isolation for closed-loop discrete-time systems; iv) fault tolerant control design for nonlinear discrete-time systems; and v) detection of simultaneous faults may occur at the same time.

References

[1] J. Chen, R. J. Patton. *Robust Model-based Fault Diagnosis for Dynamic Systems*. Boston: Kluwer, 1999.

[2] S. Simani, C. Fantuzzi, R. J. Patton. *Model-based Fault Diagnosis in Dynamic Systems Using Identification Techniques*. London: Springer, 2003.

[3] H. Hammouri, M. Kinnaert, E. H. Yaagoubi. Observer-based approach to fault detection and isolation for nonlinear systems. *IEEE Transactions on Automatic Control*, 1999, 44(10): 1879 – 1884.

[4] C. Edwards, S. K. Spurgeon, R. J. Patton. Sliding mode observers for fault detection and isolation. *Automatica*, 2000, 36(4): 541 – 553.

[5] C. D. Persis, A. Isidori. A geometric approach to nonlinear fault detection and isolation. *IEEE Transactions on Automatic Control*, 2001, 46(6): 853 – 865.

[6] B. Jiang, M. Staroswiecki, V. Cocquempot. Fault diagnosis based on adaptive observer for a class of nonlinear systems with unknown parameters. *International Journal of Control*, 2004, 77(4): 415 – 426.

[7] F. Caccavale, P. Cilibrizzi, F. Pierri, et al. Actuators fault diagnosis for robot manipulators with uncertain model. *Control Engineering Practice*, 2009, 17(1): 146 – 157.

[8] Q. Shen, B. Jiang, V. Cocquempot. Fault tolerant control for T-S fuzzy systems with application to near space hypersonic vehicle with actuator faults. *IEEE Transaction on Fuzzy Systems*, 2012, 20(4): 652 – 665.

[9] L. Yao, J. Qin, H. Wang, et al. Design of new fault diagnosis and fault tolerant control scheme for non-Gaussian singular stochastic distribution systems. *Automatica*, 2012, 48(9): 2305 – 2313.

[10] H. Wang, Z. Huang, S. Daley. On the use of adaptive updating rules for actuator and sensor fault diagnosis. *Automatica*, 1997, 33(2): 217 – 225.

[11] M. M. Polycarpou, P. A. Ioannou. Modelling, identification and stable adaptive control of continuous-time nonlinear dynamical systems using neural networks. *Proceedings of the American Control Conference*, Evanston: American Automatic Control Council, 1992: 36 – 40.

[12] M. A. Demetriou, M. M. Polycarpou. Incipient fault diagnosis of dynamical systems using online approximators. *IEEE Transactions on Automatic Control*, 1998, 43(11): 1612 – 1617.

[13] M. M. Polycarpou, A. B. Trunov. Learning approach to nonlinear fault diagnosis: detectability analysis. *IEEE Transactions on Automatic Control*, 2000, 45(4): 806 – 812.

[14] X. Zhang, M. M. Polycarpou, T. Parisini. A robust detection and isolation scheme for abrupt and incipient faults in nonlinear systems. *IEEE Transactions on Automatic Control*, 2002, 47(4): 576 – 593.

[15] Z. Mao, B. Jiang, P. Shi. Fault-tolerant control for a class of nonlinear sampled-data systems via a Euler approximate observer. *Automatica*, 2010, 46(11): 1852 – 1859.

[16] F. Caccavale, F. Pierri, L. Villani. Adaptive observer for fault diagnosis in nonlinear discrete-time systems. *Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME*, 2008, 130(2): DOI 10.1115/1.2837310.

[17] T. T. Balaje, S. Jagannathan. An online approximator-based fault detection framework for nonlinear discrete-time systems. *Proceedings of the 46th IEEE Conference on Decision and Control*, Piscataway: IEEE, 2007: 2608 – 2613.

[18] T. T. Balaje, S. Jagannathan. A model-based fault-detection and prediction scheme for nonlinear multivariable discrete-time systems with asymptotic stability guarantees. *IEEE Transactions on Neural Networks*, 2010, 21(3): 404 – 423.

[19] M. G. Riccardo, G. Ferrari, T. Parisini, et al. Distributed fault detection and isolation of large-scale discrete-time nonlinear systems: an adaptive approximation approach. *IEEE Transactions on Automatic Control*, 2012, 57(2): 275 – 290.

[20] C. Wang, D. J. Hill. Learning from neural control. *IEEE Transactions on Neural Networks*, 2006, 17(1): 30 – 46.

[21] C. Wang, D. J. Hill. Deterministic learning and rapid dynamical pattern recognition. *IEEE Transactions on Neural Networks*, 2007, 18(3): 617 – 630.

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

[23] C. Wang, T. Chen. Rapid detection of small oscillation faults via deterministic learning. *IEEE Transactions on Neural Networks*, 2011, 22(8): 1284 – 1296.

[24] C. Wang, B. Wen, W. Si, et al. Modeling and detection of rotating stall in axial flow compressors – Part I: Investigation on high-order M-G models via deterministic learning. *Acta Automatica Sinica*, 2014, 40(7): 1265 – 1277.

[25] W. Zeng, C. Wang. Gait recognition across different walking speeds via deterministic learning. *Neurocomputing*, 2015, 152: 139 – 150.

[26] W. Zeng, C. Wang. Classification of neurodegenerative diseases using gait dynamics via deterministic learning. *Information Sciences*, 2015, 317: 246 – 258.

[27] X. Dong, C. Wang, J Hu. Electrocardiogram (ECG) pattern modeling and recognition via deterministic learning. *Control Theory and Technology*, 2014, 12(4): 333 – 344.

[28] E. W. Bai, S. S. Sastry. Persistency of excitation, sufficient richness and parameter convergence in discrete time adaptive control. *Systems & Control Letters*, 1985, 6(3): 153 – 163.

[29] T. Liu, C. Wang, D. J. Hill. Deterministic learning and rapid dynamical pattern recognition of discrete-time systems. *Proceedings of the IEEE International Symposium on Intelligent Control*, Piscataway: IEEE, 2008: 1091 – 1096.

[30] C. Wang, T. R. Chen, T. Liu. Deterministic learning and data-based modeling and control. *Acta Automatica Sinica* (*Chinese*), 2009, 35(6): 693 – 706.

[31] C. Yuan, C. Wang. Design and performance analysis of deterministic learning of sampled-data nonlinear systems. *Science China Information Sciences*, 2014, 57(18): DOI 10.1007/s11432-012-4731-3.

[32] L. P. Shilnikov, A. L. Shilnikov, D. V. Turaev, et al. *Methods of Qualitative Theory in Nonlinear Dynamics – Part II*. Singapore: World Scientific, 2001.

[33] C. Wang, T. Chen, G. Chen, et al. Deterministic learning of nonlinear dynamical systems. *International Journal of Bifurcation and Chaos*, 2009, 19(4): 1307 – 1328.

[34] V. H. Garnier, A. H. Epstein, E. M. Greitzer. Rotating wave as a stall inception indiccation in axial compressors. *ASME Journal of Turbomachinery*, 1991, 113(2): 1 – 9.

[35] J. H. Deane, D. C. Hamill. Instability, subharmonics, and chaos in power electronic systems. *IEEE Transactions on Power Electronics*, 1990, 5(3): 260 – 268.

[36] S. R. Samantaray, P. K. Dash, S. K. Upadhyay. Adaptive Kalman fiter and neural network based high impedance fault detection in power distribution networks. *Electrical Power and Energy Systems*, 2009, 31(3): 167 – 172.

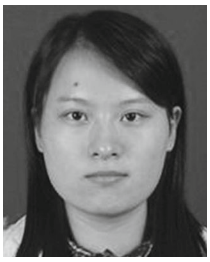
[37] A. R. Messina. *Inter-area Oscillations in Power Systems*. New York: Springer, 2009.

[38] H. Yin, P. Wang, T. Alpcan, et al. Hopf bifurcation and oscillations in a communication network with heterogeneous delays. *Automatica*, 2009, 45(10): 2358 – 2367.

[39] H. G. Hosseini, D. Luo, K. J. Reynolds. The comparison of different feed forward neural network architectures for ECG signal diagnosis. *Medical Engineering & Physics*, 2006, 28(4): 372 – 378.

[40] C. Yuan, C. Wang. Pesistency of excitation and performance of deterministic learning. *Systems & Control Letters*, 2011, 60(12): 952 – 959.

[41] U. Altinisik, M. Yildirim. A new fault tolerant control approach for the three-tank system using data mining. *Computers & Electrical Engineering*, 2012, 38(6): 1627 – 1635.



Junmin HU is a Ph.D. candidate at the Center for Control and Optimization, School of Automation Science and Engineering, South China University of Technology. Her research interest covers adaptive NN control/identification, deterministic learning theory and oscillation fault diagnosis. E-mail: hujunmin852@163.com.

Cong WANG received the B.E. and M.E. degrees from Beijing University of Aeronautic & Astronautics in 1989 and 1997, respectively, and the Ph.D. degree from the Department of Electrical & Computer Engineering, The National University of Singapore in 2002. From 2001 to 2004, he did his postdoctoral research at the Department of Electronic Engineering, City University of Hong Kong. He has been with the School of Automation Science and Engineering, South China University of Technology, Guangzhou, China, since 2004, where he is currently a professor. He has authored and co-authored over 60 international journal and conference papers and the book Deterministic Learning Theory for Identification, Recognition and Control. He serves as an Associate Editor of the IEEE Transactions on Neural Networks and Learning Systems since 2012, and as an Associate Editor for Control Theory and Technology, and ACTA AUTOMATICA SINICA (two best journals in systems and control area in China) since 2008 and 2011, respectively. He is a member of the Technical Committee on Intelligent Control of the IEEE CSS. His research interest includes intelligent control, neural networks, nonlinear systems and control, dynamical pattern recognition, pattern-based control, dynamical systems, and oscillation fault diagnosis. E-mail: wangcong@scut.edu.cn.

Xunde DONG is a Ph.D. candidate at the Center for Control and Optimization, School of Automation Science and Engineering, South China University of Technology. His research interest covers adaptive NN control/identification, deterministic learning theory and distributed parameter system. E-mail: dong.xd@mail.scut.edu.cn.