ELSEVIER

Journal of Process Control 13 (2003) 407–421

JOURNAL OF PROCESS CONTROL

www.elsevier.com/locate/jprocont

Subspace identification of continuous time models for process fault detection and isolation☆

Weihua Li[1], Harigopal Raghavan, Sirish Shah*

Department of Chemical and Materials Engineering, University of Alberta, Edmonton, AB, T6G 2G6, Canada

Received 22 August 2001; received in revised form 2 April 2002; accepted 26 June 2002

Abstract

This paper proposes a novel subspace approach towards identification of optimal residual models for process fault detection and isolation (PFDI) in a multivariate continuous-time system. We formulate the problem in terms of the state space model of the continuous-time system. The motivation for such a formulation is that the fault gain matrix, which links the process faults to the state variables of the system under consideration, is always available no matter how the faults vary with time. However, in the discrete-time state space model, the fault gain matrix is only available when the faults follow some known function of time within each sampling interval. To isolate faults, the fault gain matrix is essential. We develop subspace algorithms in the continuous-time domain to directly identify the residual models from sampled noisy data without separate identification of the system matrices. Furthermore, the proposed approach can also be extended towards the identification of the system matrices if they are needed. The newly proposed approach is applied to a simulated four-tank system, where a small leak from any tank is successfully detected and isolated. To make a comparison, we also apply the discrete time residual models to the tank system for detection and isolation of leaks. It is demonstrated that the continuous-time PFDI approach is practical and has better performance than the discrete-time PFDI approach.

Keywords: Subspace methods of identification; Numerical integration; Process fault detection and isolation; Multivariate continuous-time systems

1. Introduction

This paper proposes a novel subspace approach towards identification of optimal residual models for *process fault detection and isolation* (PFDI) in a multivariate continuous-time (CT) system. We formulate the problem in terms of the CT state space model, i.e. {**A**, **B**, **C**, **D**} of the system, because most physical systems are CT by nature. Representation of CT systems by discrete-time (DT) models is only an approximation of their dynamics. Since it is easier to identify DT models than CT models and in many cases a CT system's dynamics can be well represented by a DT model, DT models have been widely used with success. However, in some cases, e.g. in the isolation of process faults, one has to use CT models. For example, while detection and isolation of sensor faults in a CT system can be performed using its DT state space model [5,12,21], most of the work done on PFDI depends on the CT state space model [7,13,15,27] of the system.

Note that to perform PFDI, besides knowing the normal state space model: {**A**, **B**, **C**, **D**} the fault gain matrix that links the *process faults* to the state variables of the system is indispensable. We show in this paper that for the CT state space model of the system, the fault gain matrix is always available no matter how the faults are varying with time. However, in the DT state space model of the system, the fault gain matrix is available only when the faults follow some known functions of time, e.g. being piece-wise constant within one sampling interval. Without knowing the fault gain matrix, while process fault detection can still be carried out, process fault isolation is absolutely impossible.

Most of the well known PFDI approaches, e.g. the observer-based approaches and the Chow–Willsky approach [6,7,13,15,27,29] assumes the availability of a CT state space model of the system considered. In a

☆ The earlier version of this paper was presented at the IFAC sponsored CHEMFAS-4 meeting in June, 2001, Korea.

* Corresponding author. Tel.: +1-780-492-5162; fax: +1-780-492-2881.

E-mail address: sirish.shah@ualberta.ca (S. Shah).

[1] Present address: Matrikon Inc., Suite 1800, 10405 Jasper Avenue, Edmonton, AB, T5J 3N4, Canada.

0959-1524/03/$ - see front matter
PII: S0959-1524(02)00066-5

limited number of cases, a CT state space model of the system can be obtained from first principles when the mechanisms of the system are well understood. However, in complex systems such as chemical engineering processes, building the CT state space models by means of the first principles is extremely difficult. An empirical process model by use of subspace methods of identification (SMI) is a good alternative.

SMI has become an active research area and has been successfully applied to identification of multivariate DT state space models since the late 1980s [17–20,23,24]. In comparison with the traditional prediction errors methods (PEM) of identification [14] SMI has better numerical properties for systems with high dimensionality. For instance, a detailed remark regarding the advantage of SMI over the PEM can be seen in [25]. SMI has also been applied to the identification of primary residual models (PRM) for sensor fault detection and isolation [12,21].

Recently, SMI for CT models has been proposed [11]. However, this approach has a drawback: it is sensitive to the initial values of CT signals. In this paper, we choose the numerical integrators proposed by [22] to transform the signals and their derivatives in the CT domain into sampled data. We propose a SMI and apply it to the transformed signals for identification of the primary residual model (PRM) for process fault detection without identifying the system matrices: $\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$, explicitly. The chosen numerical integrators are immune to initial values of the CT signals, computationally efficient, and have sufficient accuracy. Furthermore, we transform the PRM into a set of structured residual models (SRMs) for fault isolation. The SRMs are designed such that each structured residual vector (SRV) generated by one SRM is most sensitive to only one single fault while insensitive to the other faults.

The newly proposed approach is applied to a simulated tank system, where detection and isolation of leaks in any tanks is successfully conducted. Moreover, although this approach is oriented for residual models identification, it can be easily extended to the identification of the complete system model $\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$, if necessary. For example, in many control relevant problems, a knowledge of $\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$ is usually required.

This paper is organized as follows. Section 2 is devoted to motivation and problem formulation. Section 3 outlines the numerical integrators to be used to transform the CT signals into DT data. Identification of the PRM for fault detection is given in Section 4. In Section 5, identification of a set of SRMs for isolation of single and multiple faults is investigated. The proposed approach is numerically evaluated in Section 6, where its effectiveness in detection and isolation of leaks in a simulated tank system is demonstrated. In addition, to make a comparison, the detection and isolation of leaks by means of the DT residual models has also been performed. The paper ends in Section 7 with conclusions and remarks. A supplementary result concerned with the estimation of $\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$ is given in the Appendices.

2. Motivation and problem formulation

2.1. Motivation

As depicted in Fig. 1, a four-tank system with leak $\delta_i(t)$ in the ith tank, $\forall i \in [1,4]$ can be represented by the following equations [10]:

$$\begin{aligned}
\frac{dh_1}{dt} &= \frac{\gamma_1 k_1}{A_1}\upsilon_1 - \frac{a_1}{A_1}\sqrt{2gh_1} + \frac{a_3}{A_1}\sqrt{2gh_3} - \frac{\epsilon_1\delta_1}{A_1}\\
\frac{dh_2}{dt} &= \frac{\gamma_2 k_2}{A_2}\upsilon_2 - \frac{a_2}{A_2}\sqrt{2gh_2} + \frac{a_4}{A_4}\sqrt{2gh_2} - \frac{\epsilon_2\delta_2}{A_2}\\
\frac{dh_3}{dt} &= \frac{(1-\gamma_2)k_2}{A_3}\upsilon_2 - \frac{a_3}{A_3}\sqrt{2gh_3} - \frac{\epsilon_3\delta_3}{A_3}\\
\frac{dh_4}{dt} &= \frac{(1-\gamma_1)k_1}{A_4}\upsilon_1 - \frac{a_4}{A_4}\sqrt{2gh_4} - \frac{\epsilon_4\delta_4}{A_4}
\end{aligned} \tag{1}$$

where, A_i is the cross-section of tank i, a_i is the cross-section of the ith outlet hole, and h_i is the water level in tank i. Furthermore, in Eq. (1), υ_i is the voltage applied to Pump i, and $k_i\upsilon_i$ is the corresponding flow. The parameters $\gamma_1, \gamma_2 \in (0, 1)$ are determined from how the valves are set prior to an experiment. The flow to tank 1 is $\gamma_1 k_1 \upsilon_1$, the flow to tank 4 is $(1-\gamma_1)k_1\upsilon_1$ and similarly for tanks 2 and 3. The acceleration due to gravity is denoted by g. The measured level signals are $k_c h_1, k_c h_2, k_c h_3$ and $k_c h_4$ where k_c is a parameter associated with the sensor gain. In addition,

$$\epsilon_i = \begin{cases} 1, \text{if the } i\text{th tank leaks.} \\ 0, \text{if the } i\text{th does not leak.} \end{cases}$$

Eq. (1) can be linearized around an operating point ([10]): $\{h_1^o, h_2^o, h_3^o, h_4^o, \upsilon_1^o, \upsilon_2^o\}$, where h_i^o and, $\upsilon_j^o, \forall i \in [1,4]$

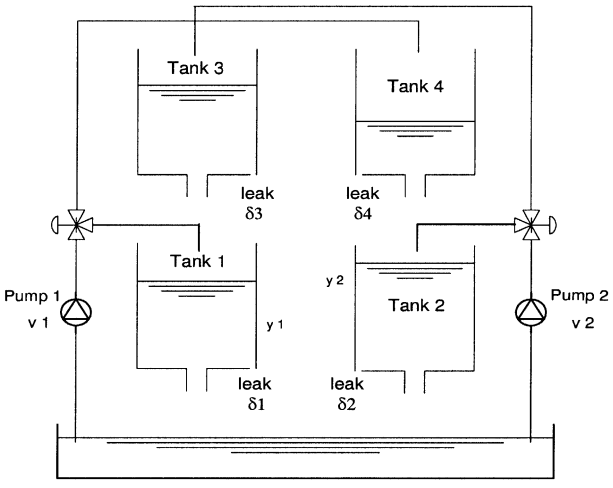

Fig. 1. Schematic of the four water tank system.

and $j \in [1,2]$, are the values of levels and voltages, respectively. As a consequence, the following state-space model can be obtained:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{\Xi}_{d_i}\mathbf{f}_{d_i}(t)$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) \tag{2}$$

where, $\mathbf{x}(t) = \left[h_1^-(t)\ h_2^-(t)\ h_3^-(t)\ h_4^-(t)\right]^T$ is the state variable vector in terms of *deviation variables* with $h_i^-(t) = h_i(t) - h_i^o, \forall i = 1,4$, $\mathbf{u}(t) = \left[\upsilon_1(t) - \upsilon_1^o\ \ \upsilon_2(t) - \upsilon_2^o\right]^T$ is the input vector, $\mathbf{y}(t) = \left[k_c h_1^-(t)\ k_c h_2^-(t)\ k_c h_3^-(t)\ k_c h_4^-(t)\right]^T$ is the output vector, and $\{\mathbf{A}, \mathbf{B}, \mathbf{C}\}$ are system matrices, assuming that some or all the elements of $\mathbf{x}(t)$ are measurable. Details of $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ can be seen in [10]. Further, in Eq. (2), $\mathbf{f}_{d_i}(t) \in \Re^{d_i}$ is the equivalent fault magnitude vector containing elements of $\left[\frac{-\delta_1^-(t)}{A_1} \cdots \frac{-\delta_4^-(t)}{A_4}\right]^T$ with $\mathbf{\Xi}_{d_i}$ standing for the associated columns of the 4×4 identity matrix $\mathbf{I}_4, \forall d_i \in [1,4]$. Note that $\delta_i^-(t) = \delta_i(t) - \delta_i^o$, where δ_i^o is the value of $\delta_i(t)$ at an operating point, $\forall i \in [1,4]$.

Although Eq. (2) is derived for the four-tank system, it can be generalized to describe a wide class of multivariate linear time-invariant CT systems with process faults as follows:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{\Xi}_{d_i}\mathbf{f}_{d_i}(t)$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \tag{3}$$

where $\mathbf{u}(t) \in \Re^q$, $\mathbf{y}(t) \in \Re^m$, $\mathbf{x}(t) \in \Re^n$, $\mathbf{f}_{d_i}(t) \in \Re^{d_i}$, $\mathbf{\Xi}_{d_i} \in \Re^{n \times d_i}$, $\forall d_i \in [1,n]$, and $\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$ are similar to those defined in Eq. (2) (therein $\mathbf{D} = \mathbf{0}$) with compatible dimensions. We assume that $\{\mathbf{C}, \mathbf{A}\}$ is observable. Furthermore, in the sequel throughout the paper, we assume that the order n is known, because schemes for the order determination in the subspace identification framework are available [1]. Note that in Eq. (3), the fault model is simply $\mathbf{\Xi}_{d_i}$, i.e. columns of the identity matrix $\mathbf{I}_n \in \Re^{n \times n}$. However, as will be shown in the next paragraph, the fault model in the DT state space model is not always available.

Discretizing Eq. (3) with a sampling period T_s leads to the following DT state space model [2]

$$\mathbf{x}(k+1) = \mathbf{A}_d\mathbf{x}(k) + \mathbf{B}_d\mathbf{u}(k) + \mathbf{e}^{\mathbf{A}T_s}\int_0^{T_s}\mathbf{e}^{-\mathbf{A}\tau}\mathbf{\Xi}_{d_i}\mathrm{f}_{d_i}(kT_s + \tau)\mathrm{d}\tau$$
$$\mathbf{y}(k) = \mathbf{C}_d\mathbf{x}(k) + \mathbf{D}_d\mathbf{u}(k) \tag{4}$$

where

$$\mathbf{A}_d = \mathbf{e}^{\mathbf{A}T_s},\ \mathbf{B}_d = \int_0^{T_s}\mathbf{e}^{\mathbf{A}\tau}\mathbf{B}\mathrm{d}\tau$$
$$\mathbf{C}_d = \mathbf{C},\ \mathbf{D}_d = \mathbf{D} \tag{5}$$

and the input is assumed to be invariant within each sampling interval, i.e. $\mathbf{u}(kT_s + \tau) = \mathbf{u}(k)$ for $0 < \tau \leqslant \tau_s$.

In Eq. (4), if $\mathbf{f}_{d_i}(kT_s + \tau)$ is assumed to be invariant within a sampling interval, $\mathbf{f}_{d_i}(kT_s + \tau) = \mathbf{f}_{d_i}(kT_s), \forall \tau \in (0, T_s)$, the fault related term is then equal to

$$\left(\mathbf{e}^{\mathbf{A}T_s}\int_0^{T_s}\mathbf{e}^{-A\tau}\mathbf{\Xi}_{d_i}\mathrm{d}\tau\right)\mathbf{f}_{d_i}(k),$$

where we denote $\mathbf{f}_{d_i}(kT_s) = \mathbf{f}_{d_i}(k)$. Consequently, the fault gain matrix in the DT state space equation is

$$\mathbf{e}^{\mathbf{A}T_s}\int_0^{T_s}\mathbf{e}^{-\mathbf{A}\tau}\mathbf{\Xi}_{d_i}d\tau = \mathbf{A}_d\mathbf{A}^{-1}\left(\mathbf{I}_n - \mathbf{A}_d^{-1}\right)\mathbf{\Xi}_{d_i}.$$

However, almost invariably this assumption will not hold, and therefore the fault-contribution term in the right hand side (RHS) of Eq. (4) is

$$\mathbf{e}^{\mathbf{A}T_s}\int_0^{T_s}\mathbf{e}^{-\mathbf{A}\tau}\mathbf{\Xi}_{d_i}\mathbf{f}_{d_i}(kT_s + \tau)\mathrm{d}\tau$$

In this case, since $\mathbf{f}_{d_i}(kT_s + \tau)$ is time varying, $\forall \tau \in [0, T_s]$, the fault gain matrix in the DT state space is not available. Without such a matrix, while fault detection can still be carried out, nothing can be done with respect to fault isolation in terms of the Chow–Willsky scheme [5] and any observer-based scheme [6]. In particular, the assumption can be severely misleading if the sampling period T_s is relatively large. To obtain the aforementioned fault gain matrix, one way is to increase the sampling frequency $\frac{1}{T_s}$ significantly. However, this can increase the cost of collecting data and cause numerical issues in system identification and control as pointed out by [16]. Therefore, for the purpose of process fault isolation in a CT system, a knowledge of its CT state space representation is essential. Note that even if the assumption of piece-wise constancy holds, a knowledge of $\mathbf{A}$ in the CT state space equation is necessary for the calculation of the DT fault gain matrix: $\mathbf{A}_d\mathbf{A}^{-1}\left(\mathbf{I}_n - \mathbf{A}_d^{-1}\right)\mathbf{\Xi}_{d_i}$

2.2. Problem formulation

Differentiating Eq. (3) repeatedly, we obtain the ith derivative of $\mathbf{y}(t)$ as follows:

$$\mathbf{y}^{(i)}(t) = \mathbf{C}\mathbf{A}^i\mathbf{x}(t) + \mathbf{C}\sum_{j=0}^{i-1}\mathbf{A}^{i-1-j}\left[\mathbf{B}\mathbf{u}^{(j)}(t) + \mathbf{\Xi}_{d_i}\mathbf{f}_{d_i}^{(j)}(t)\right] + \mathbf{D}\mathbf{u}^{(i)}(t),$$

where $i \in [1,s]$, and $\left\{\mathbf{u}^{(j)}(t), \mathbf{f}_{d_i}^{(j)}(t)\right\}$ are the jth derivatives of $\mathbf{u}(t)$ and $\mathbf{f}_{d_i}(t)$, respectively. Subsequently, by stacking, we arrive at

$$\mathbf{y}_s(t) = \mathbf{\Gamma}_s\mathbf{x}(t) + \mathbf{H}_s\mathbf{u}_s(t) + \mathbf{G}_s\mathbf{\Xi}_{s,d_i}\mathbf{f}_{s,d_i}(t) \tag{6}$$

where, $\Xi_{s,d_i} = \mathbf{I}_{s+1} \otimes \Xi_{d_i} \in \Re^{n_{s\times d_i(s+1)}}$, and $\otimes$ denotes the Kronecker tensor product;

$$\mathbf{y}_s(t) = \begin{bmatrix} \mathbf{y}(t) \\ \mathbf{y}^{(1)}(t) \\ \vdots \\ \mathbf{y}^{(s)}(t) \end{bmatrix} \in \Re^{m_s}$$

is the stacked output vector;

$$\mathbf{\Gamma}_s = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{\mathrm{s}} \end{bmatrix} \in \Re^{m_s\times n}$$

is the extended observability matrix; and

$$\mathbf{H}_s = \begin{bmatrix} \mathbf{D} & 0 & \dots & 0 \\ \mathbf{CB} & \mathbf{D} & & \vdots \\ \vdots & \ddots & \ddots & \\ \mathbf{CA}^{S\text{-}1} & \mathbf{B}\dots & \mathbf{CB} & \mathbf{D} \end{bmatrix} \in \Re^{m_s\times q_s}$$

is a lower triangular block Toeplitz matrix. In addition,

$$\mathbf{G}_s = \begin{bmatrix} 0 & \dots & & 0 \\ \mathbf{C} & 0 & & \vdots \\ \vdots & \ddots & \ddots & \\ \mathbf{CA}^{s-1} & \dots & \mathbf{C} & 0 \end{bmatrix} \in \Re^{m_s\times n_s}$$

is analogous to $\mathbf{H}_s$ in structure. Note that $\mathbf{G}_s$ is completely dependent on the first ms rows of $\mathbf{\Gamma}_s$. Therefore, once $\mathbf{\Gamma}_s$ is identified, from it $\mathbf{G}_s$ can be derived. In addition, in Eq. (6), s is defined as the order of the parity space or the maximum detectability index of the fault [13] and is selected to be equal to n throughout the paper without loss of generality; $m_s=(s+1)m$; $n_s=(s+1)n$; and $q_s=(s+1)q$. The stacked vectors $\mathbf{u}_s(t) \in \Re^{q_s}$ and $\mathbf{f}_{s,d_i}(t)$ are expressed in the same formats as $\mathbf{y}_s(t)$.

We introduce

$$\widetilde{\mathbf{H}}_s = [\mathbf{I}_{\mathrm{ms}}| - \mathbf{H}_s] \in \Re^{m_s\times(m_s+q_s)}$$

where $\mathbf{I}_{ms} \in \Re^{m_s\times m_s}$ is an identity matrix, then we can rewrite Eq. (6) as

$$\widetilde{\mathbf{H}}_s \begin{bmatrix} \mathbf{y}_s(t) \\ \mathbf{u}_s(t) \end{bmatrix} = \mathbf{\Gamma}_s\mathbf{x}(t) + \mathbf{G}_s\Xi_{s,d_i}\,\mathbf{f}_{s,d_i}(t) \tag{7}$$

Pre-multiplying Eq. (7) by a matrix $\mathbf{W}_0$, which lies in the null space of $\mathbf{\Gamma}_s$, i.e. $\mathbf{W}_0\mathbf{\Gamma}_s = 0$, produces

$$\epsilon_s^c(t) \equiv \mathbf{P}_s \begin{bmatrix} \mathbf{y}_s(t) \\ \mathbf{u}_s(t) \end{bmatrix} = \mathbf{W}_0\mathbf{G}_s\Xi_{\ s,d_i}\,\mathbf{f}_{s,d_i}(t) \in \Re^{sn} \tag{8}$$

where $\mathbf{P}_s \equiv \mathbf{W}_0\widetilde{\mathbf{H}}_s = [\mathbf{W}_0| - \mathbf{W}_0\mathbf{H}_s]$ is defined as the PRM for fault detection. Note that $\mathbf{W}_0\mathbf{G}_s$ is the fault model in $\epsilon_s^c(t)$.

By extending the Chow–Willsky scheme in the DT domain [5] to the CT domain, we define $\epsilon_s^c(t)$ as the continuous-time primary residual vector (CT-PRV), because it is zero if no fault occurs, i.e. $\mathbf{f}_{d_i}(t) = 0$; or nonzero if a fault occurs, i.e. $\mathbf{f}_{d_i}(t) \neq 0$. In order to ensure that the process fault $\mathbf{f}_{d_i}(t)$ is detected with the highest sensitivity, we design $\mathbf{W}_0$ to have maximum covariance with $\mathbf{G}_s$ while it is orthogonal to $\mathbf{\Gamma}_s$. Mathematically we establish the following objective functions, $\forall i,j = \{[1, \mathrm{Rank}(\mathbf{G}_s)]\} \bigcap \{i \neq j\}$.

$$J_0^i = \max_{\mathbf{W}_0(i,:)} \mathbf{W}_0(i,:)\mathbf{G}_s\mathbf{G}_s^T\mathbf{W}_0^T(i,:) - \mathbf{W}_0(i,:)\mathbf{\Gamma}_s\lambda_0^i - \lambda_1^i\left(\|\mathbf{W}_0(i,:)\|_2 - 1\right) - \lambda_2^i\mathbf{W}_0(i,:)\mathbf{W}_0^T(j,:) \tag{9}$$

where $\mathbf{W}_0(i;:)$ and $\mathbf{W}_0(j;:)$ are the ith and jth rows of $\mathbf{W}_0$, respectively; $\{\lambda_0^i \in \Re^n, \lambda_1^i, \lambda_2^i\}$ are the Lagrange multipliers; and $||\ ||_2$ is the 2-norm of a vector. Note that in the objective functions, $\mathbf{W}_0(i,:)\mathbf{\Gamma}_s = 0$ and the orthonormality among the rows of $\mathbf{W}_0$ have been taken into consideration.

In parallel with the development in [12], the optimal solution to $\mathbf{W}_0$ can be obtained as follows:

$\mathbf{W}_0^T$ = Eigenvectors related to non-zero eigenvalues of the matrix $\mathbf{\Gamma}_s^{\perp}\mathbf{G}_s\mathbf{G}_s^T$,

where $\mathbf{\Gamma}_s^{\perp} = \mathbf{I}_{m_s} - \mathbf{\Gamma}_s\mathbf{\Gamma}_s^{+}$, and $^{+}$ indicates the Moore–Penrose pseudo inverse of the argument. Note that $\mathbf{\Gamma}_s^{\perp}\mathbf{G}_s\mathbf{G}_s^T$ is a non-symmetric matrix. To calculate $\mathbf{W}_0$, we first translate this non-symmetric eigenproblem into a symmetric eigenproblem. Then any existing standard algorithms can be used. The calculation of $\mathbf{W}_0$ is given in Appendix. Since $\mathbf{\Gamma}_s$ has rank n, $\mathbf{\Gamma}_s^{\perp}$ will have rank m_s-n. Furthermore, note that $\mathbf{G}_s\mathbf{G}_s^T$ has a rank equal to $\mathrm{Rank}(\mathbf{G}_s)$. As a result, $\mathbf{W}_0$ will have min $\{m_s$–n, $\mathrm{Rank}(\mathbf{G}_s)\}$ non-zero eigenvalues. Assume that $m \geqslant n$ and $\mathrm{Rank}(\mathbf{C}) = n$, we can then conclude min $\{m_s$–n, $\mathrm{Rank}(\mathbf{G}_s)\} = \mathrm{Rank}(\mathbf{G}_s) = sn$, which is the number of independent rows in $\mathbf{W}_0$, i.e. $\mathbf{W}_0 \in \Re^{sn\times m_s}$.

Assume that both the sampled inputs $\mathbf{u}^{\mathrm{o}}(k)$ and outputs $\mathbf{y}^{\mathrm{o}}(k)$ are corrupted by independently Gaussian distributed white noise vectors $\mathbf{v}(k) \in \Re^q$ and $\mathbf{o}(k) \in \Re^m$, which have covariance matrices $\mathbf{R}_v$ and $\mathbf{R}_o$, respectively, i.e.

$$\mathbf{u}^{\mathrm{o}}(k) = \mathbf{u}(k) + \mathbf{v}(k),\ \mathbf{y}^{\mathrm{o}}(k) = \mathbf{y}(k) + \mathbf{o}(k) \tag{10}$$

This is referred to as the errors-in-variables (EIV) case ([4]). Note that the two noise vectors are mutually independent and are independent of $\mathbf{u}(k)$ and $\mathbf{y}(k)$. The problem of residual model identification then can be stated as follows:

Given the sampled inputs and outputs when $\mathbf{f}_{\mathrm{di}}(t)=\mathbf{0}$, *identify the PRM*: $\mathbf{P}_s$, *which includes the consistent identification of* $\mathbf{\Gamma}_s$, *calculation of* $\mathbf{W}_0$, *and identification of* $\mathbf{W}_0\mathbf{H}_s$. *Furthermore, from the PRM design a set of SRMs for fault isolation.*

3. Overview of the numerical integrators

To identify $\mathbf{\Gamma}_s$ and $\mathbf{W}_0\mathbf{H}_s$, computationally we need $\mathbf{u}_s(t)$ and $\mathbf{y}_s(t)$, which contain the CT signals $\{\mathbf{u}(t), \mathbf{y}(t)\}$ and their derivatives from the 1st up to the sth order. The derivatives are not directly measurable and many approaches have been developed to deal with them [3,28]. Herein we use the numerical integrator proposed in [22] to transform the derivatives into discrete data, because such an integrator has a number of attractive features: simplicity of implementation, insensitivity to initial values of the CT signals, and high accuracy.

The integral of a CT signal, e.g. $\mathbf{u}(t)$, over the time interval $[t-lT_s, t]$ can be approximately calculated by

$$\varsigma_1[\mathbf{u}(t)] = \int_{t-lT_s}^{t}\mathbf{u}(\tau)\mathrm{d}\tau \approx \omega_0\mathbf{u}(t)+\ldots+\omega_l\mathbf{u}(t-lT_s) = \sum_{i=0}^{l}\omega_i q^{-i}\mathbf{u}(t) \tag{11}$$

where T_s is the integration step size, chosen to be the same as the sampling interval for easy implementation; l is considered as the length factor of the integrator (a natural number); and q^{-1} is the unit delay operator, i.e. $q^{-1}\mathbf{u}(t)=\mathbf{u}(t-T_s)$. The filter coefficients ω_i depend on the type of numerical integration methods to be employed. For instance, when the trapezoidal integration rule is used, the filter coefficients are [22]:

$$\omega_0=\omega_l=\frac{T_s}{2},\ \omega_i=T_s,\ i\in[1,l-1]$$

Similarly, the sth multiple integral of the ith derivative $\mathbf{u}^{(i)}(t)$ of $\mathbf{u}(t)$ can be defined as follows:

$$\varsigma_s\left[\mathbf{u}^{(i)}(t)\right] = \int_{t-lT_s}^{t}\int_{\tau_1-lT_s}^{\tau_1}\cdots\int_{\tau_{s-1}-lT_s}^{\tau_{s-1}}\mathbf{u}^{(i)}(\tau)\mathrm{d}\tau_s\ldots\mathrm{d}\tau_1,\quad i\in[0,s]. \tag{12}$$

Furthermore, Eq. (12) can be approximately calculated by

$$\varsigma_s\left[\mathbf{u}^{(i)}(t)\right]=\phi_{s,i}\left(q^{-1}\right)\mathbf{u}(t),\ i\in[0,s] \tag{13}$$

where,

$$\phi_{s,i}\left(q^{-1}\right)=\left(1-q^{-l}\right)^i\left(\omega_0+\omega_1q^{-1}+\cdots+\omega_lq^{-l}\right)^{s-i}=\sum_{\upsilon=0}^{sl}\beta_\upsilon^i q^{-\upsilon}$$

and β_υ^i is the coefficient in the polynomial $\phi_{s,i}\left(q^{-1}\right)$. A rigorous proof of Eq. (13) and comments on the numerical integrator can be found in [22]. Note that the optimal choice of the filter length is an open issue especially for multivariate systems. The approach of minimizing the prediction error proposed by [26] might provide a solution.

4. Identification of the PRM

4.1. Descriptions of the identification model

When $\mathbf{f}_{s,d_i}(t)=0$, Eq. (6) is reduced to

$$\mathbf{y}_s(t)=\mathbf{\Gamma}_s\mathbf{x}(t)+\mathbf{H}_s\mathbf{u}_s(t) \tag{14}$$

Performing multiple integration of Eq. (14) s times by using Eq. (13), we obtain

$$\phi_s\left[q^{-1},\mathbf{y}(t)\right]=\mathbf{\Gamma}_s\phi_{s,0}\left(q^{-1}\right)\mathbf{x}(t)+\mathbf{H}_s\phi_s\left[q^{-1},\mathbf{u}(t)\right]+\mathbf{e}_s(t,T_s) \tag{15}$$

where, for any vector $\xi(t)$, e.g. $\xi(t)=\mathbf{u}(t)$ or $\xi(t)=\boldsymbol{y}(t)$

$$\phi_s\left[q^{-1},\xi(t)\right]=\begin{bmatrix}\phi_{s,0}\left(q^{-1}\right)\xi(t)\\ \phi_{s,1}\left(q^{-1}\right)\xi(t)\\ \vdots\\ \phi_{s,s}\left(q^{-1}\right)\xi(t)\end{bmatrix} \tag{16}$$

and such a convention will be used throughout the paper. In addition, $\mathbf{e}_s(t,T_s)$ is a truncation-error due to numerical integration of inputs and outputs.

With available sampled data $\left\{\mathbf{u}^{\mathrm{o}}(k),\mathbf{y}^{\mathrm{o}}(k)\right\}$ substituting Eq. (10) into Eq. (15) at $t=kT_s$ gives

$$\phi_s\left[q^{-1},\mathbf{y}^{\mathrm{o}}(k)\right]=\mathbf{\Gamma}_s\phi_{s,0}\left(q^{-1}\right)\mathbf{x}(k)+\mathbf{H}_s\phi_s\left[q^{-1},\mathbf{u}^{\mathrm{o}}(k)\right]+\mathbf{E}_s(k) \tag{17}$$

where,

$$\mathbf{E}_s(k)=-\mathbf{H}_s\phi_s\left[q^{-1},\mathbf{v}(k)\right]+\phi_s\left[q^{-1},\mathbf{o}(k)\right]+\mathbf{e}_s(k,T_s)\in\Re^{m_s}$$

The first term in $\mathbf{E}_s(k)$ is a moving average (MA) process of $\mathbf{v}(k)$ and $\mathbf{o}(k)$. In addition, as addressed by [22], the truncation error $\mathbf{e}_s(k,T_s)$ can be controlled with the sampling interval T_s. For simplicity, in the following

analysis, it is assumed that a small T_s is used so that the influence of $\mathbf{e}_s(k, T_s)$ on the identification of $\mathbf{P}_s$ can be negligible, i.e. $\mathbf{e}_s(k, T_s) \approx 0$.

We form the following data matrix for inputs $\mathbf{u}^o(k)$

$$\mathbf{U}^o_{k,s,N} = \left[\phi_s\left(q^{-1}, \mathbf{u}^o(k)\right) \cdots \phi_s\left(q^{-1}, \mathbf{u}^o(k+N-1)\right)\right] \in \Re^{q_s \times N},$$

where N is the number of data samples in the matrix. Similarly, we form the output data matrix $\mathbf{Y}^o_{k,s,N} \in \Re^{m_s \times N}$. The use of the newly formed data matrices in Eq. (17) gives

$$\mathbf{Y}^o_{k,s,N} = \mathbf{\Gamma}_s \mathbf{X}_{k,N} + \mathbf{H}_s \mathbf{U}^o_{k,s,N} + \mathbf{E}_{k,s,N} \in \Re^{m_s \times N} \tag{18}$$

where,

$$\mathbf{X}_{k,N} = \left[\phi_{s,0}\left(q^{-1}\right)\mathbf{x}(k) \cdots \phi_{s,0}\left(q^{-1}\right)\mathbf{x}(k+N\text{-}1)\right] \in \Re^{n \times N}$$

and $\mathbf{E}_{k,s,N} \in \Re^{m_s \times N}$ resembles $\mathbf{U}^o_{k,s,N}$ or $\mathbf{Y}^o_{k,s,N}$ in format.

After the establishment of Eq. 18, the following remarks can be made:

Remark 1. The data matrices $\mathbf{U}^o_{k,s,N}$ and $\mathbf{Y}^o_{k,s,N}$ are composed of sampled inputs and outputs and their time-lagged values, respectively. Further, each data matrix has $s+1$ row blocks and N columns. For example, in the ith column of $\mathbf{U}^o_{k,s,N}$ one row block contains the linear combination of $\mathbf{u}^o(k+i-1)$ and the time-lagged values $\left\{\mathbf{u}^o(k+i-2) \cdots \mathbf{u}^o(k+i-1-sl)\right\}$ for $i \in [1, N]$.

Remark 2. The observability matrix $\mathbf{\Gamma}_s$ and the lower triangular block Topeliz matrix $\mathbf{H}_s$ in Eq. (18) are exactly the same as those in the CT model given by Eq. (14), showing that although the numerical integrator transforms the CT signals into discrete data, it preserves the original CT system model in the DT domain. This excels other CT identification approaches [3,11], which transform the CT signals and result in a different DT model at the same time.

Remark 3. The matrix $\mathbf{X}_{k,N}$ is not measurable. However, as will be shown later one does not have to know it in order to identify the models in which we are interested.

Remark 4. With a similar structure to $\mathbf{U}^o_{k,s,N}$ or $\mathbf{Y}^o_{k,s,N}$, $\mathbf{E}_{k,s,N} \in \Re^{m_s \times N}$ has $s+1$ row blocks and N columns. To obtain consistent estimates of $\mathbf{\Gamma}_s$ and $\mathbf{H}_s$, one has to remove the effect of $\mathbf{E}_{k,s,N}$ on the identification.

Remark 5. With the identified $\mathbf{P}_s$, one can generate the primary residual vector (PRV) using sampled data for fault detection. From Eqs. (6) and (17), the PRV is

$$\varepsilon_s(k) = \mathbf{P}_s \begin{bmatrix} \phi_s\left(q^{-1}, \mathbf{y}^o(k)\right) \\ \phi_{\mathbf{s}}\left(q^{-1}, \mathbf{u}^o(k)\right) \end{bmatrix} = \begin{cases} \mathbf{W}_0 \mathbf{E}_s(k), & \text{no fault} \\ \mathbf{W}_0 \mathbf{E}_s(k) + \mathbf{W}_0 \mathbf{G}_s \mathbf{\Xi}_{s,d_i} \xi_s\left[\mathbf{f}_{s,d_i}(k)\right], & \text{with fault.} \end{cases}$$

where on the RHS, (i) the first and the second lines are the computational and the internal forms of the PRV, respectively; and (ii) $\xi_s\left[\mathbf{f}_{s,d_i}(k)\right]$ is the sth multiple integral of $\mathbf{f}_{s,d_i}(t)$ with t ranging from kT_s -lT_s to kT_s, i.e.

$$\xi_s\left[\mathbf{f}_{s,d_i}(k)\right] = \int_{kT_s - lT_s}^{kT_s} \int_{\tau_1 - lT_s}^{\tau_1} \cdots \int_{\tau_{s-1} - lT_s}^{\tau_{s-1}} \mathbf{f}_{s,d_i}(\tau_s) \mathrm{d}\tau_s \ldots \mathrm{d}\tau_1$$

Remark 6. We denote the fault-free value of $\varepsilon_s(k)$ and the fault-contributed term by $\varepsilon^*_s(k) \equiv \mathbf{W}_0 \mathbf{E}_s(k)$ and $\varepsilon^f_s(k) \equiv \mathbf{W}_0 \mathbf{G}_s \mathbf{\Xi}_{s,d_i} \xi_s\left[\mathbf{f}_{s,d_i}(k)\right]$, respectively.

Apparently $\varepsilon^*_s(k)$ is also a MA process of the noise vector $\left[\mathbf{v}^T(k)\ \mathbf{o}^T(k)\right]^T$, and follows a zero mean multivariate Gaussian distribution, due to the assumed distribution of $\left[\mathbf{v}^T(k)\ \mathbf{o}^T(k)\right]^T$. Consequently, it turns out from Eq. (19) that

$$\varepsilon_s(k) = \begin{cases} \varepsilon^*_s(k) \sim \mathcal{N}\left(\mathbf{0}, \mathbf{R}_{\varepsilon,s}\right), \text{no fault} \\ \varepsilon^*_s(k) + \varepsilon^f_s(k) \sim \mathcal{N}\left(\varepsilon^f_s(k), \mathbf{R}_{\varepsilon,s}\right), \text{ with fault} \end{cases}$$

where $\mathbf{R}_{\varepsilon,s}$ is the covariance of $\varepsilon^*_s(k)$. From Eq. (17),

$$\mathbf{R}_{\varepsilon,s} = \mathbf{W}_0\left\{\mathbf{H}_s \mathbf{Cov}\left[\phi_s\left(q^{-1}\right)\mathbf{v}(k)\right]\mathbf{H}^T_s + \mathbf{Cov}\left[\phi_s\left(q^{-1}\right)\mathbf{o}(k)\right]\right\}\mathbf{W}^T_0$$

where, **Cov**( ) is the covariance of the argument. Therefore, one can perform fault detection by simply checking if $\varepsilon_s(k)$ is zero mean. We define a chi-square distributed variable $\eta_{s,o}(k) = \varepsilon^T_s(k)\mathbf{R}^{-1}_{\varepsilon,s}\varepsilon_s(k)$. With a preselected level of significance α, e.g. $\alpha = 0.01$, $\eta_{s,o}(k) > \chi^2_\alpha(ns)$ indicates that the faults have occurred.

Remark 7. The fault gain matrix in $\varepsilon_s(k)$ is $\mathbf{W}_0 \mathbf{G}_s \mathbf{\Xi}_{s,d_i}$ depending on which we can design a set of SRMs for fault isolation as will be shown in Section 5.

4.2. Consistent estimation of $\mathbf{\Gamma}_s$

With a choice of $k = 1$ in Eq. (18), we obtain

$$\mathbf{Y}^o_{1,s,N} = \mathbf{\Gamma}_s \mathbf{X}_{1,N} + \mathbf{H}_s \mathbf{U}^o_{1,s,N} + \mathbf{E}_{1,s,N} \tag{20}$$

Post-multiplying Eq. (20) by $\frac{1}{N}\left[\mathbf{U}^{oT}_{L,s,N}\ \mathbf{Y}^{oT}_{Ls,N}\right]$ gives

$$\frac{1}{N}\mathbf{Y}^o_{1,s,N}\left[\mathbf{U}^{oT}_{L,s,N}\ \mathbf{Y}^{oT}_{L,s,N}\right] = \frac{1}{N}\mathbf{\Gamma}_s \mathbf{X}_{1,N}\left[\mathbf{U}^{oT}_{L,s,N}\ \mathbf{Y}^{oT}_{L,s,N}\right] + \frac{1}{N}\mathbf{H}_s\ \mathbf{U}^o_{1,s,N}\left[\mathbf{U}^{oT}_{L,s,N}\ \mathbf{Y}^{oT}_{L,s,N}\right] + \frac{1}{N}\mathbf{E}_{1,s,N}\left[\mathbf{U}^{oT}_{L,s,N}\ \mathbf{Y}^{oT}_{L,s,N}\right] \tag{21}$$

Since the numerical integrator is an (sl)th order filter, we can show that selecting $L = sl + 2$ in Eq. (21) makes the last term on the RHS asymptotically vanish as $N \to \infty$. (19)

Substituting the following QR decomposition

$$\begin{bmatrix} \mathbf{U}^{\circ}_{1,s,N} \\ \mathbf{Y}^{\circ}_{1,s,N} \end{bmatrix} \begin{bmatrix} \mathbf{U}^{\circ T}_{L,sN} & \mathbf{Y}^{\circ T}_{L,s,N} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{11} & \mathbf{0} \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{Q}_1 \\ \mathbf{Q}_2 \end{bmatrix}$$

into Eq. (21) leads to

$$\begin{aligned} &\left[\mathbf{I}_{m_s}| - \mathbf{H}_s\right] \begin{bmatrix} \mathbf{R}_{21}\mathbf{Q}_1 + \mathbf{R}_{22}\mathbf{Q}_2 \\ \mathbf{R}_{11}\mathbf{Q}_1 \end{bmatrix} \\ &\quad = \mathbf{\Gamma}_s \mathbf{X}_{1,N} \left[\mathbf{U}^{oT}_{L,s,N}\ \mathbf{Y}^{oT}_{L,s,N}\right] \end{aligned} \tag{22}$$

where

$$\mathbf{Q}_i^T \mathbf{Q}_j = \begin{cases} \mathbf{I}, i = j, i, j \in [1, 2] \\ \mathbf{0}, i \neq j, i, j \in [1, 2] \end{cases}$$

has been employed with $\mathbf{I}$ standing for an identity matrix of appropriate dimensions.

Post-multiplying Eq. (22) by $\mathbf{Q}_2^T$ results in

$$\mathbf{R}_{22} = \mathbf{\Gamma}_s \mathbf{X}_{1,N} \left[\mathbf{U}^{\circ T}_{L,s,N}\ \mathbf{Y}^{\circ T}_{L,s,N}\right] \mathbf{Q}_2^T$$

Applying singular value decomposition (SVD) to $\mathbf{R}_{22}$, we obtain

$$\mathbf{R}_{22} = \mathbf{U}_l \mathbf{\Lambda} \mathbf{V}_r^T$$

Therefore, the first n vectors of $\mathbf{U}_l$ give consistent estimate of $\mathbf{\Gamma}_s$ (up to a column space), i.e.

$$\mathbf{\Gamma}_s = \mathbf{U}_l(:, 1 : n) \tag{23}$$

assuming that $\mathbf{X}_{1,N}\left[\mathbf{U}^{\circ T}_{L,s,N}\ \mathbf{Y}^{\circ T}_{L,s,N}\right]\mathbf{Q}_2^T$ has rank n.

To ensure the validity of this assumption, we have to fully understand the consistency conditions. At this moment, we concentrate on the identification of $\mathbf{P}_s$. However, the consistency analysis is under further investigation.

4.3. Identification of $\mathbf{W}_0\mathbf{H}_s$

With the identified $\mathbf{\Gamma}_s$, one can derive $\mathbf{G}_s$ and calculate $\mathbf{W}_0$ following the steps shown earlier. Further, pre-multiplying Eq. (22) by $\mathbf{W}_0$ yields

$$[\mathbf{W}_0| - \mathbf{W}_0\mathbf{H}_s] \begin{bmatrix} \mathbf{R}_{21}\mathbf{Q}_1 + \mathbf{R}_{22}\mathbf{Q}_2 \\ \mathbf{R}_{11}\mathbf{Q}_1 \end{bmatrix} = \mathbf{0}$$

where $\mathbf{W}_0\mathbf{\Gamma}_s \equiv \mathbf{0}$ is employed. Post-multiplying the preceding equation by $\mathbf{Q}_1^T$ gives

$$\mathbf{W}_0\mathbf{R}_{21} = \mathbf{W}_0\mathbf{H}_s\mathbf{R}_{11}$$

Consequently

$$\mathbf{W}_0\mathbf{H}_s = \mathbf{W}_0\mathbf{R}_{21}\mathbf{R}_{11}^{+} \tag{24}$$

Eventually, $\mathbf{P}_s = \left[\mathbf{W}_0| - \mathbf{W}_0\mathbf{R}_{21}\mathbf{R}_{11}^{+}\right]$

5. Design of the SRMs for fault isolation

With the identified $\mathbf{P}_s$, we can generate the PRV for fault detection. Further, to isolate faults, we can transform the PRV into a set of structured residual vectors (SRVs) [12], where one SRV is designed to be immune to a specified subset of faults, but has maximized sensitivity to the remaining faults. The performance of the vector-based fault isolation is much better than the performance of the scalar residual based isolation, as has been proved in [12]. Here we extend the SRVs to the isolation of process faults in the CT systems. To generate SRVs, we can design a set of SRMs, including the selection of an incidence matrix to characterize the SRVs. The incidence matrix is dependent on the number of faults to be isolated, the system order n and the order of the parity space, and is not unique. A detailed discussion with respect to the selection can be found in [8,9,12].

In the system under consideration given in Eq. (2), the dimension of the fault magnitude vector $\mathbf{f}_{d_i}(t)$ can be up to n, i.e. $1 \leqslant d_i \leqslant n$. In the simplest case, we only have to isolate a single fault. However, in the most difficult case, we have to isolate up to n multiple faults, although the probability for multiple faults to occur simultaneously is small. Hence there are $\sum_{i=0}^{n} C_i^n = \sum_{i=0}^{n} \frac{n!}{i!(n-i)!}$ fault scenarios in total, where C_i^n is the combination of i from n and $i!$ is the factorial. To isolate these scenarios, an ideal design is to transform the PRV into n SRVs, where the ith SRV is affected with highest sensitivity only by the ith fault, while it is immune to all the other faults, $\forall i \in [1, n]$. As will be analyzed later, such an ideal design can be achievable provided that certain conditions are satisfied.

Computationally, the ith SRV is equal to

$$\begin{aligned} \mathbf{r}_{s,i}(k) = \mathbf{W}_i \mathbf{\varepsilon}_s(k) &= \mathbf{W}_i \mathbf{P}_s \begin{bmatrix} \phi_s\left(q^{-1}, \mathbf{y}^{\circ}(k)\right) \\ \phi_s\left(q^{-1}, \mathbf{u}^{\circ}(k)\right) \end{bmatrix} \\ &= \mathbf{W}_i\left[\mathbf{\varepsilon}_s^*(k) + \mathbf{\varepsilon}_s^f(k)\right] \end{aligned} \tag{25}$$

where $\mathbf{W}_i\mathbf{P}_s$ is defined as the ith SRM. Denote $\mathbf{W}_0\mathbf{G}_s \in \Re^{(sn)\times n_s}$ by $\tilde{\mathbf{G}}_s$, since $\mathbf{\varepsilon}_s^f(k) = \mathbf{W}_0\mathbf{G}_s\mathbf{\Xi}_{s,d_i}\zeta_s\left[\mathbf{f}_{s,d_i}(k)\right]$, we have the fault-contributed term:

$$r_{s,i}^f(k) = \mathbf{W}_i\mathbf{\varepsilon}_s^f(k) = \mathbf{W}_i\tilde{\mathbf{G}}_s\mathbf{\Xi}_{s,d_i}\zeta_s\left[\mathbf{f}_{s,d_i}(k)\right]$$

in $\mathbf{r}_{s,i}(k)$. Note that $\tilde{\mathbf{G}}_s$ has ns non-zero columns because its last n columns are zeros. Apparently, if $\mathbf{r}_{s,i}(k)$ is designed to be insensitive to all but the ith fault $f_i(t)$, $\mathbf{W}_i$ should be orthogonal to the $(n-1)s$ columns associated with

$$[f_1(t)..f_{i-1}(t)\ f_{i+1}(t)..f_n(t)]$$

in $\tilde{\mathbf{G}}_s$, $\forall i \in [1, n]$.

Denote

$$\tilde{\mathbf{G}}_{s,i} \equiv \left[\tilde{\mathbf{G}}_s(:, i)|\tilde{\mathbf{G}}_s(:, n+i), ..., \tilde{\mathbf{G}}_s(:, (s-1)n+i)\right] \in \Re^{(sn)\times s}, \forall i = 1, n$$

and the remaining columns in $\tilde{\mathbf{G}}_s$ after leaving out $\tilde{\mathbf{G}}_{s,i}$ by $\tilde{\mathbf{G}}^-_{s,i} \in \Re^{(sn)\times((n-1)s)}$, respectively. It is desired that $\mathbf{W}_i$ has maximized covariance with $\tilde{\mathbf{G}}_s$ under the constraint $\mathbf{W}_i\tilde{\mathbf{G}}^-_{s,i}$. Since $\tilde{\mathbf{G}}^-_{s,i} \in \Re^{(sn)\times((n-1)s)}$, $\mathbf{W}_i$ will have $sn-(n-1)s = s$ independent rows. Similarly, it can be shown that, $\forall i \in [1, n]$:

$$\mathbf{W}_i^T = \text{Eigenvectors related to non-zero eigen values of } \tilde{\mathbf{G}}^{-,\perp}_{s,i}\tilde{\mathbf{G}}_s\tilde{\mathbf{G}}_s^T$$

where, $\tilde{\mathbf{G}}^{-,\perp}_{s,i} = \mathbf{I}_{sn} - \tilde{\mathbf{G}}^-_{s,i}\left(\tilde{\mathbf{G}}^-_{s,i}\right)^+ \in \Re^{(sn)\times(sn)}$.

Eq. (25) can be further expressed as

$$\mathbf{r}_{s,i}(k) = \mathbf{r}^*_{s,i}(k) + \mathbf{r}^f_{s,i}(k) \in \Re^n \tag{26}$$

where $\mathbf{r}^*_{s,i} = \mathbf{W}_i\varepsilon^*_s(k)$ follows a zero mean Gaussian distribution with covariance $\mathbf{R}^i_{\varepsilon,s} = \mathbf{W}_i\mathbf{R}_{\varepsilon,s}\mathbf{W}_i^T$. Hence $f_i(t) \neq 0$ can be detected and uniquely isolated if the mean of $\mathbf{r}_{s,i}(k)$ is non-zero. Moreover, we define a chi-square distributed *isolation index*

$$\eta_{s,i}(k) = \mathbf{r}^T_{s,i}(k)\left(\mathbf{R}^i_{\varepsilon,s}\right)^{-1}\mathbf{r}_{s,i}(k),$$

which has a threshold $\chi^2_\alpha(n)$. We denote the normalized isolation index by $\bar{\eta}_{s,i}(k) \equiv \eta_{s,i}(k)/\chi^2_\alpha(n)$. Then $\bar{\eta}_{s,i}(k) > 1$ indicates the presence of $|f_i(t)| \neq 0$ (the mean of $\mathbf{r}_{s,i}(k)$ is non-zero), while $\bar{\eta}_{s,i}(k) < 1$ indicates $|f_i(t)| = 0$ (the mean of $\mathbf{r}_{s,i}(k)$ is zero). In addition, to isolate multiple faults, e.g. $\{f_i(t) \neq 0\} \bigcap \{f_j(t) \neq 0\}$, we can simply check if $\bar{\eta}_{s,i}(k)$ and $\bar{\eta}_{s,j}(k)$ are greater than 1, simultaneously.

For easy reference, we give the incidence matrix in Table 1 to describe how the n SRVs are correlated with the n faults $[f_1(t) \ \ldots \ f_n(t)]$, where a "0" means the insensitivity of one SRV to a fault (the corresponding isolation index is smaller than the threshold), while a "1" indicates that an SRV has highest sensitivity to a fault (the corresponding isolation index is larger than the threshold).

Eventually, the complete procedure of identifying the PRM and SRMs, as well as performing PFDI in a CT system using sampled data, is summarized as follows:

(1) Given a set of training data, construct $\{\mathbf{Y}^\circ_{1,s,N}, \mathbf{U}^\circ_{1,s,N}, \mathbf{Y}^\circ_{L,s,N}, \mathbf{U}^\circ_{L,s,N}, \}$ with $L = sl+2$.
(2) Identify $\mathbf{\Gamma}_s$ using Eq. (23). Derive $\mathbf{G}_s$ from $\mathbf{\Gamma}_s$ and calculate $\mathbf{W}_0$. Identify $\mathbf{W}_0\mathbf{H}_s$ using Eq. (24). Then obtain $\mathbf{P}_s = [\mathbf{W}_0 | -\mathbf{W}_0\mathbf{H}_s]$.
(3) Select an incidence matrix as illustrated in Table 1 to characterize the SRVs. Subsequently, calculate $\mathbf{W}_i, \forall i \in [1, n]$. Consequently, the ith SRM is $\mathbf{W}_i\,\mathbf{P}_s$.
(4) With newly sampled test data, first construct

$$\{\phi_s[q^{-1}, \mathbf{u}^\circ(k)], \phi_s[q^{-1}, \mathbf{y}^\circ(k)]\}$$

as shown in Eq. (16). Then, calculate the PRV as shown in the first line of Eq. (19) for fault detection, and a set of SRVs as shown in Eq. (25) for fault isolation.

6. Numerical results

In this section, we use the water tank system [10] illustrated in Fig. 1 as a pilot plant to demonstrate the effectiveness of the proposed approach for the identification of residual models. Further, to make a comparison we also use DT residual models to carry out detection and isolation of leaks in the tanks. As will be shown, the CT residual models are much more powerful than their DT counterparts at isolating fast time-varying faults.

The parameter values and the chosen operating point of the laboratory process are given in Table 2. Consequently, linearizing Eq. (1) under these conditions gives the following CT state space equations [10]:

Table 1
Incidence matrix to characterize the isolation logic

	$f_1(t)$	$f_2(t)$	$f_3(t)$	...	$f_n(t)$
$\mathbf{r}_{s,1}(k)\left(\bar{\eta}_{s,1}(k)\right)$	1	0	0	...	0
$\mathbf{r}_{s,2}(k)\left(\bar{\eta}_{s,2}(k)\right)$	0	1	0	...	0
$\mathbf{r}_{s,3}(k)\left(\bar{\eta}_{s,3}(k)\right)$	0	0	1	⋱	0
⋮	⋮	⋱	⋱	⋱	⋮
$\mathbf{r}_{s,n}(k)\left(\bar{\eta}_{s,n}(k)\right)$	0	...	...	...	1

Table 2
Parameter values and chosen operating point of the laboratory process

Parameters	Units	Values
A_1; A_3	(cm^2)	28
A_2; A_4	(cm^2)	32
a_1; a_3	(cm^2)	0.071
a_2; a_4	(cm^2)	0.057
k_1; k_2	(cm^3/Vs)	3.33, 3.35
h°_1, h°_2	(cm)	12.4, 12.7
h°_3, h°_4	(cm)	1.8, 1.4
v°_1, v°_2	(V)	3.00
k_c	(V/cm)	0.50
g	(cm/s^2)	981
γ_1; γ_2		0.70, 0.60

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \begin{bmatrix} -0.0159 & 0 & 0.0419 & 0 \\ 0 & -0.0111 & 0 & 0.0333 \\ 0 & 0 & -0.0419 & 0 \\ 0 & 0 & 0 & -0.0333 \end{bmatrix} \mathbf{x}(t) \\ &+ \begin{bmatrix} 0.0833 & 0 \\ 0 & 0.0628 \\ 0 & 0.0479 \\ 0.0312 & 0 \end{bmatrix} \mathbf{u}(t) \\ &- \begin{bmatrix} \frac{1}{28} & 0 & 0 & 0 \\ 0 & \frac{1}{32} & 0 & 0 \\ 0 & 0 & \frac{1}{28} & 0 \\ 0 & 0 & 0 & \frac{1}{32} \end{bmatrix} \begin{bmatrix} \epsilon_1 \delta_1^-(t) \\ \epsilon_2 \delta_2^-(t) \\ \epsilon_3 \delta_3^-(t) \\ \epsilon_4 \delta_4^-(t) \end{bmatrix} \mathbf{y}(t) \\ &= \begin{bmatrix} 0.5 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0.5 \end{bmatrix} \mathbf{x}(t) \end{aligned} \tag{27}$$

The noise-free CT inputs to the system, i.e. $\mathbf{u}(t)$, are simulated by pseudo random binary signals with small magnitude. The frequency band for the frequency contents of the inputs is chosen to be [0, 0.03], expressed in fractions of the Nyquist frequency (see the *idinput* command in MATLAB).

First, we use SIMULINK in the CT domain to simulate Eq. (27) in the *fault-free case*, i.e. $\epsilon_i = 0$ for $i \in [1, 4]$, generating the *noise-free* and *fault-free CT outputs*. Subsequently, the CT input and output signals are sampled with a sampling interval T_s, giving $\{\mathbf{u}(k), \mathbf{y}(k)\}$. Further, an independently Gaussian distributed white noise vector with covariance $0.1^2\,\mathbf{I}_6$, i.e. $\mathcal{N}(\mathbf{0}, 0.1^2\,\mathbf{I}_6)$, in the DT domain is introduced to the *sampled noise-free inputs and outputs*, resulting in a set of noise-contaminated training data with 5000 samples.

To identify the CT residual models, two important parameters: the sampling time T_s and the length parameter of the integral filter l need to be determined.

The effect of T_s and l on the estimation of the CT model parameters is shown in Fig. 2, where the estimation error is defined by:

$$e_{est} = ||eig(A_{est}) - eig(A)||_2$$

with *eig*( ) standing for a vector containing all the eigenvalues of a matrix in a descending order. T_s and l are chosen to be equal to 2 and 14, respectively in this case.

With the chosen l and T_s, using the training data we identify $\mathbf{\Gamma}_s \in \Re^{20\times 4}$, from which we derive $\mathbf{G}_s \in \Re^{20\times 20}$. Further we calculate $\mathbf{W}_s \in \Re^{16\times 20}$ and identify $\mathbf{W}_0\mathbf{H}_s \in \Re^{16\times 10}$, resulting in $\mathbf{P}_s = [\mathbf{W}_0 | -\mathbf{W}_0\mathbf{H}_s] \in \Re^{16\times 30}$. Furthermore, in accordance with the fault isolation logic summarized in Table 1, we calculate four transformation matrices $\{\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3, \mathbf{W}_4\}$ and the corresponding SRMs: $\mathbf{W}_i\mathbf{P}_s \in \Re^{4\times 30}$, for $i \in [1, 4]$.

We use the training data and the identified $\mathbf{P}_s$ to generate a sequence of PRVs, from which we estimate the covariance $\mathbf{R}_{\varepsilon,s} \in \Re^{16\times 16}$. Moreover, we calculate covariances $\mathbf{R}^i_{\varepsilon,s} = \mathbf{W}_i\mathbf{R}^i_{\varepsilon,s}\mathbf{W}_i^T \in \Re^{4\times 4}$ of the SRVs: $\mathbf{r}_{s,i}(k)$, for $i \in [1, 4]$, respectively. Due to limited space in this paper we have not reproduced these matrices.

Faults are introduced as leaks in any tanks as mentioned earlier. It turns out from the laws of fluid mechanics that a leak is proportional to the square root of the water level in the tank, and is time-varying, i.e.

$$\delta_i^-(t) = a_i^f\sqrt{2gh_i^-(t)},\ \ \forall i \in [1, 4]$$

where, a_i^f is the size of the leak orifice. Clearly, such a fault is not piece-wise constant because $h_i^-(t)$ can vary with time within one sampling interval.

In the presence of fault(s), we also use SIMULINK to simulate Eq. (27), generating *noise-free* faulty CT signals. For instance, when we simulate a leak in tank 3, we choose $\epsilon_3 = 1$, while $\epsilon_i = 0$, for $i = 1, 2, 4$. We sample the CT signals with the same period $T_s = 2$ s. Then, we introduce a white noise vector $\mathcal{N}(\mathbf{0},0.1^2\mathbf{I}_6)$ to the sampled noise-free signals. We conduct PFDI in the following four cases:

Case 1. A leak simulated by $a_i^f = 0.01$ cm^2 is introduced to a single tank starting at the 1000th s. The relevant PFDI results are shown in Fig. 3, where the detection and isolation is done instantly without any time delay. In the figure, all the isolation indices have been scaled to have unit threshold and the x-axis represents times, i.e. $t = kT_s$, in seconds rather than number of samples, i.e. k. Note that in the following cases, all the fault isolation indices have also been scaled to have unit threshold, and the x-axis in each corresponding figure also represents times $t = kT_s$ in seconds. Since $\bar{\eta}_{s,i}(t)$ is greater than its confidence limit 1, but $\bar{\eta}_{s,i}(t)$ are all less than their confidence limits, $\forall i \in [2, 4]$, resulting in an incidence vector [1 0 0 0], it can be inferred from the isolation logic given in Table 1 that tank 1 is leaking.

Case 2. Two leaks simulated with the choice of a $a_{i_1}^f = a_{i_2}^f = 0.01^*(t - 1000)/2000$ cm^2 are introduced to two different tanks simultaneously after the 1000th s. In this case, note that the magnitude of a leak is $0.01(t-1000)/2000$, which increases very slowly. This type of fault is called the *incipient fault*, and its detection is difficult. As shown in Fig. 4, our proposed scheme is still able to detect the incipient faults after they have occurred for a period of time, e.g. when $t \geqslant 1553$ seconds. Furthermore, since $\bar{\eta}_{s,3}(t)$ and

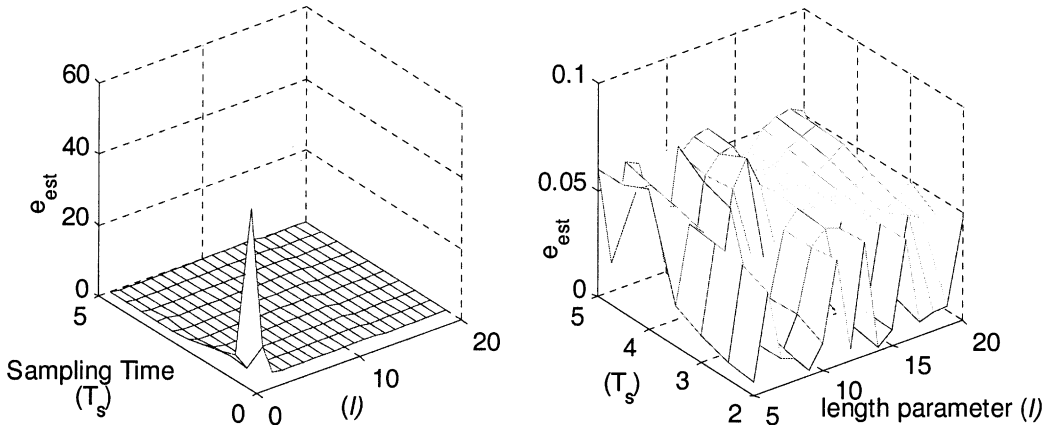

Fig. 2. Selection of T_s and l for continuous time system identification.

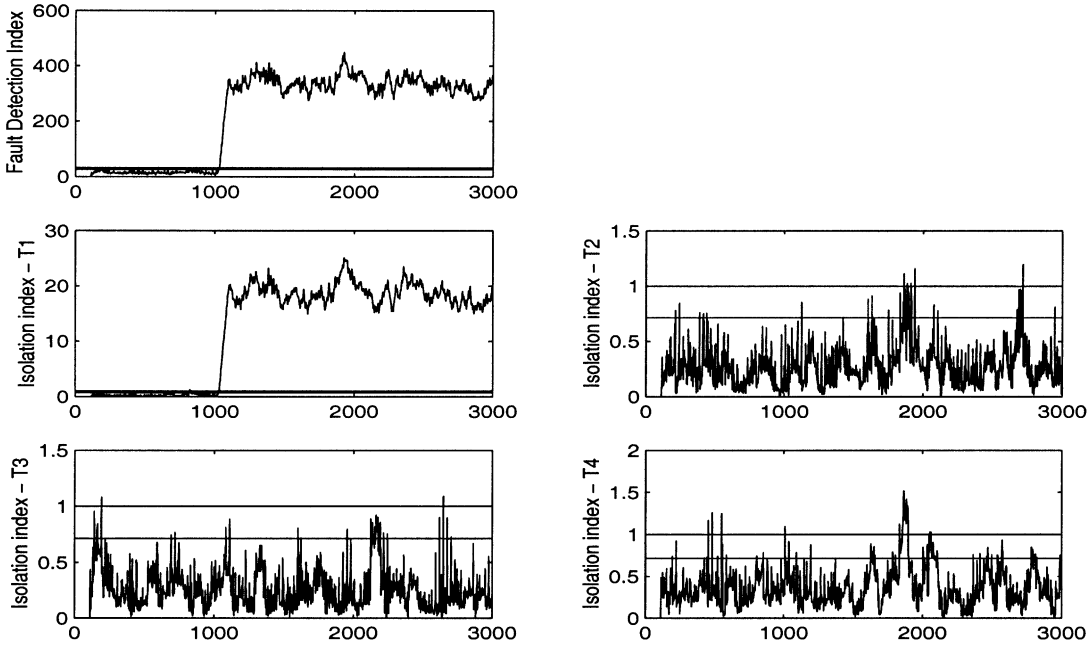

Fig. 3. CT residual model-based detection and isolation of a leak in tank 1. The incidence vector for isolation is $\left[\bar{\eta}_{s,1}(t)\ \bar{\eta}_{s,2}(t)\bar{\eta}_{s,3}(t)\ \bar{\eta}_{s,4}(t)\right]|_{t=kT_s} = \left[1\quad 0\quad 0\quad 0\right]$.

$\bar{\eta}_{s,4}(t)$ are greater than their confidence limits 1, but $\bar{\eta}_{s,1}(t)$ and $\bar{\eta}_{s,2}(t)$ are lower than their confidence limits, it is indicated that tanks 3 and 4 are leaking at the same time.

Case 3. Three leaks with a $a^f_{i_1} = a^f_{i_2} = a^f_{i_3} = 0.01$ cm, $\forall i_1$, i_2, $i_3 \in [1,4]$, are introduced to three different tanks simultaneously after the 1000th s and the PFDI results are shown in Fig. 5. After immediate detection of faults, since $\bar{\eta}_{s,2}(t)$, $\bar{\eta}_{s,3}(t)$ and $\bar{\eta}_{s,4}(t)$ are greater than their confidence limits 1 simultaneously, but $\bar{\eta}_{s,1}(t)$ is lower than the confidence limit, it can be concluded that tanks 2, 3 and 4 are leaking.

Case 4. Leaks with $a^f_{i_1} = a^f_{i_2} = a^f_{i_3} = a^f_{i_4} = 0.01$ cm^2, $\forall i_1$, i_2, i_3, $i_4 \in [1,4]$ are introduced to all four tanks simultaneously after the 1000th s. As Fig. 6 shows, the faults are detected promptly. Furthermore, since $\bar{\eta}_{s,i}(t), \forall i = 1, 2, 3, 4$, are all greater than their confidence limits 1, it is known that all the tanks are leaking.

6.1. Comparison of CT and DT residual models-based PFDI

To make a comparison, we use the DT state space model to design the PRM and SRMs. With $T_s = 2$, we obtain $\mathbf{A}_d = \mathbf{e}^{\mathbf{A}T_s}$, $\mathbf{B}_d = \int_0^{T_s} \mathbf{e}^{\mathbf{A}\tau}\mathbf{B}d\tau$, $\mathbf{C}_d = \mathbf{C}$, $\mathbf{D}_d = \mathbf{D}$, and the fault model $\mathbf{e}^{\mathbf{A}T_s}\int_0^{T_s}\mathbf{e}^{-\mathbf{A}\tau}\mathbf{\Xi}_{d_i}d\tau$ assuming piece-wise constancy of the faults, where the true values of **A**, **B**, **C**, and **D** are used. Subsequently, we compute the DT counterparts of PRM and four SRMs similarly by

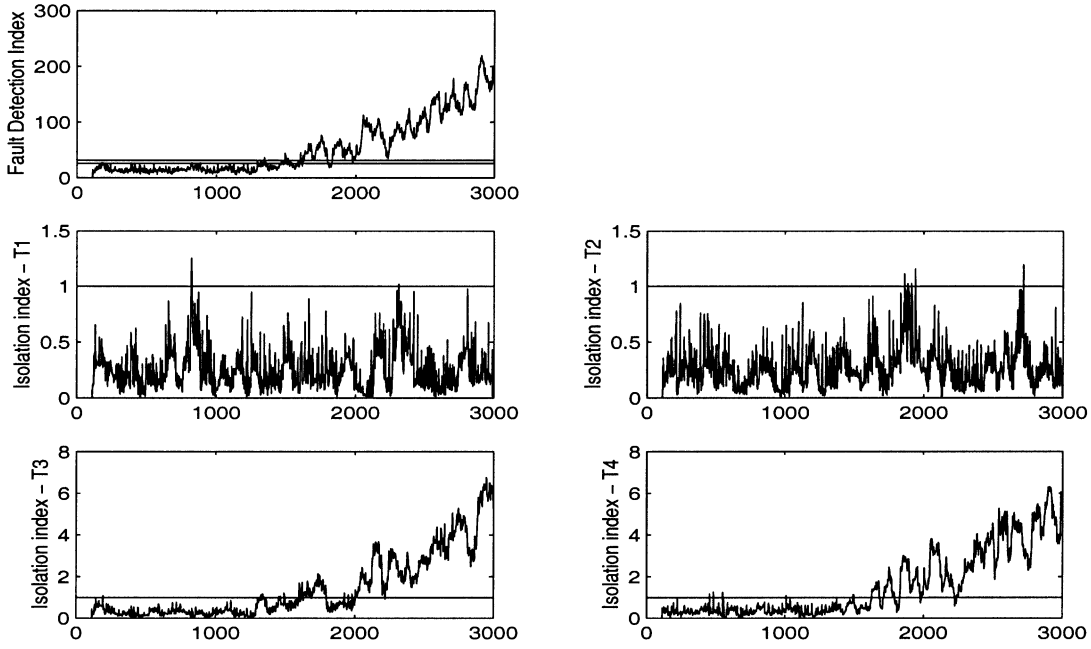

Fig. 4. CT residual model-based detection and isolation of two simultaneous leaks in tanks 3 and 4. The incidence vector for isolation is $\left[\bar{\eta}_{s,1}(t)\ \bar{\eta}_{s,2}(t)\ \bar{\eta}_{s,3}(t)\ \bar{\eta}_{s,4}(t)\right]|_{t=kT_s} = \left[0\quad 0\quad 1\quad 1\right]$.

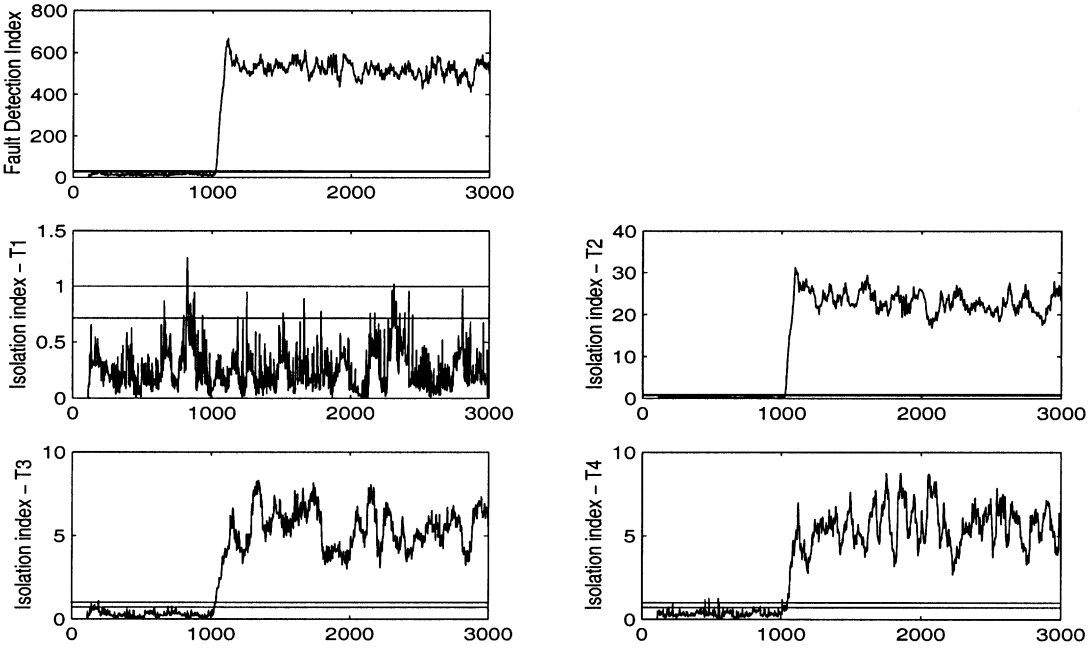

Fig. 5. CT residual model-based detection and isolation of three simultaneous leaks in tanks 2, 3 and 4. The incidence vector for isolation is $\left[\bar{\eta}_{s,1}(t)\ \bar{\eta}_{s,2}(t)\ \bar{\eta}_{s,3}(t)\ \bar{\eta}_{s,4}(t)\right]|_{t=kT_s} = \left[0\quad 1\quad 1\quad 1\right]$.

simply replacing $\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$ with $\{\mathbf{A}_d, \mathbf{B}_d, \mathbf{C}_d, \mathbf{D}_d\}$, respectively.

We introduce a single leak to a tank. Then we apply the CT and DT residual models to the same test data, respectively. It is shown in Figs. 7 and 8 that while the identified CT residual models successfully detect and isolate a leak in tank 3, the DT residual models fail to detect and isolate the time-varying fault even if they are calculated from the exactly known **A**, **B**, **C** and **D**. Eventually we introduce a ZOH to the fault, assuming that the fault is piecewise constant within one sampling interval. As depicted in Fig. 9, in this case, the

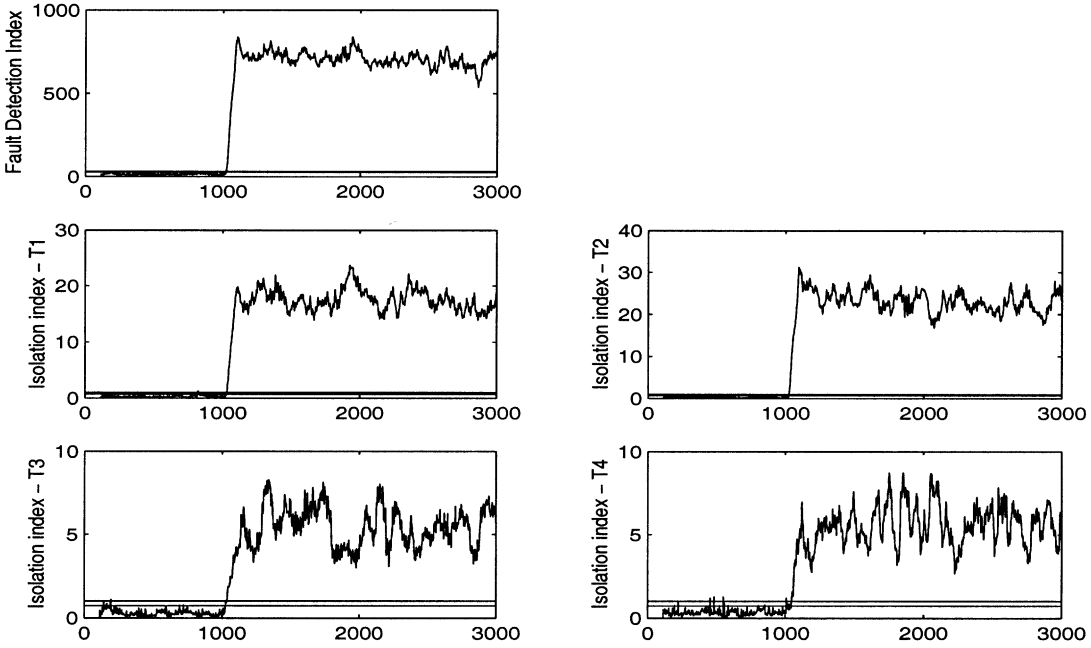

Fig. 6. CT residual model-based detection and isolation of simultaneous leaks in all 4 tanks. The incidence vector for isolation is $\left[\bar{\eta}_{s,1}(t)\ \bar{\eta}_{s,2}(t)\ \bar{\eta}_{s,3}(t)\ \bar{\eta}_{s,4}(t)\right]|_{t=kT_s} = \left[1\quad 1\quad 1\quad 1\right]$.

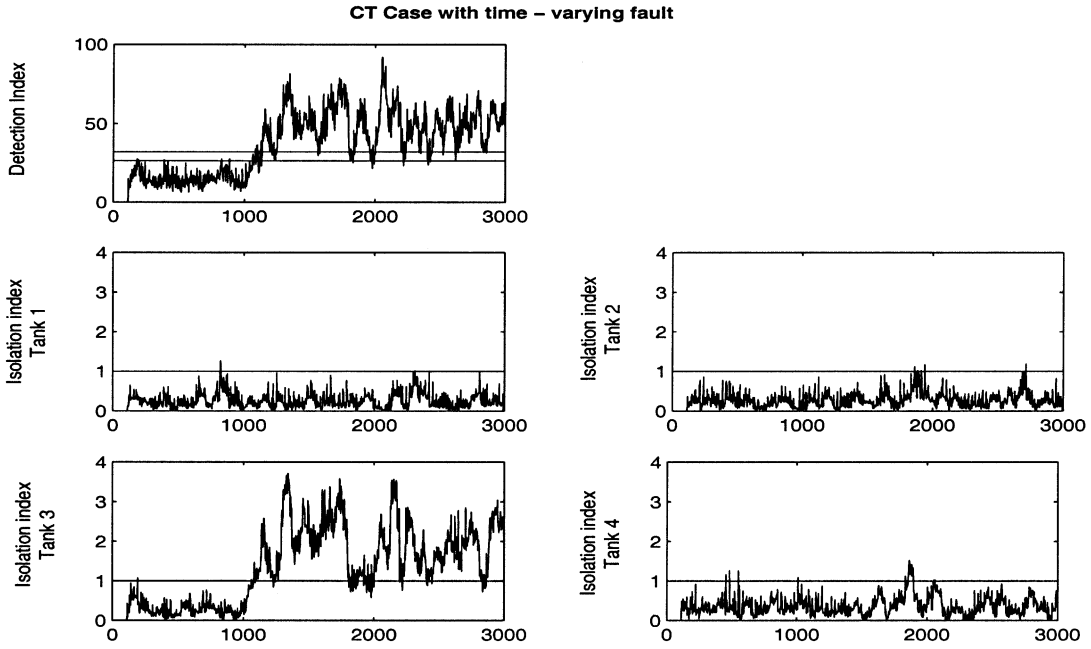

Fig. 7. The CT residual model-based PFDI scheme successfully detects and isolates a time-varying leak in tank 3. The incidence vector for isolation is $\left[\bar{\eta}_{s,1}(t)\ \bar{\eta}_{s,2}(t)\ \bar{\eta}_{s,3}(t)\ \bar{\eta}_{s,4}(t)\right]|_{t=kT_s} = \left[0\quad 0\quad 1\quad 0\right]$.

performance of fault detection and isolation based on the DT residual models is improved as compared with the performance displayed in Fig. 8. However, apparently it is still much worse than the CT residual models-based performance shown in Fig. 7.

7. Conclusions

A novel subspace approach to the identification of the PRM for process fault detection and the design of a set of SRMs for process fault isolation in multivariate CT

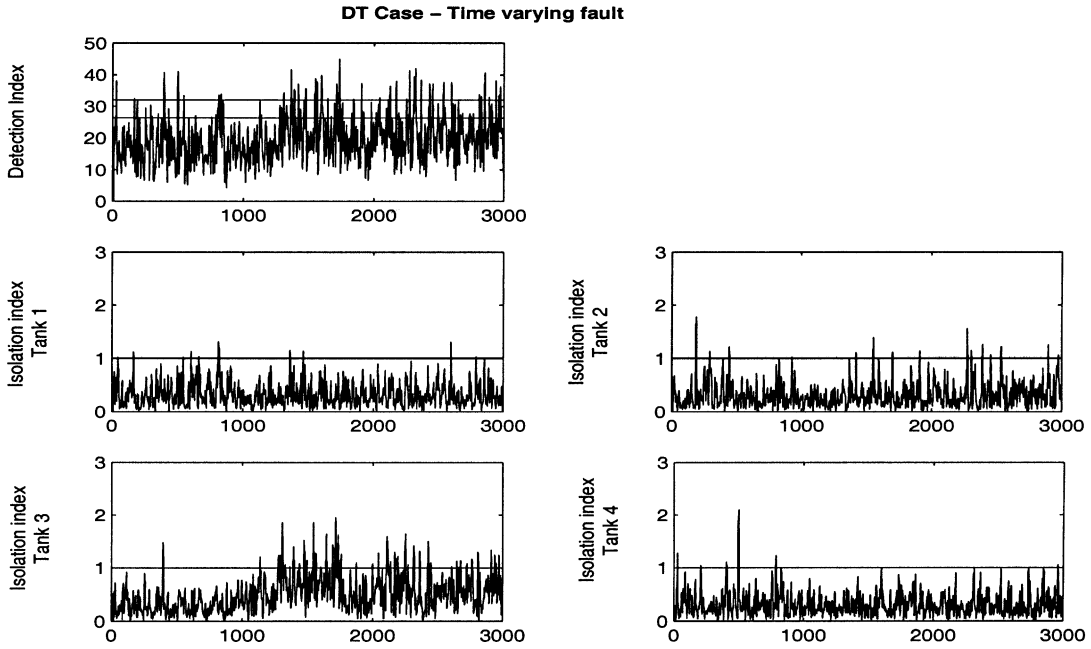

Fig. 8. The DT residual model-based PFDI scheme fails to detect and isolate a time-varying leak in tank 3.

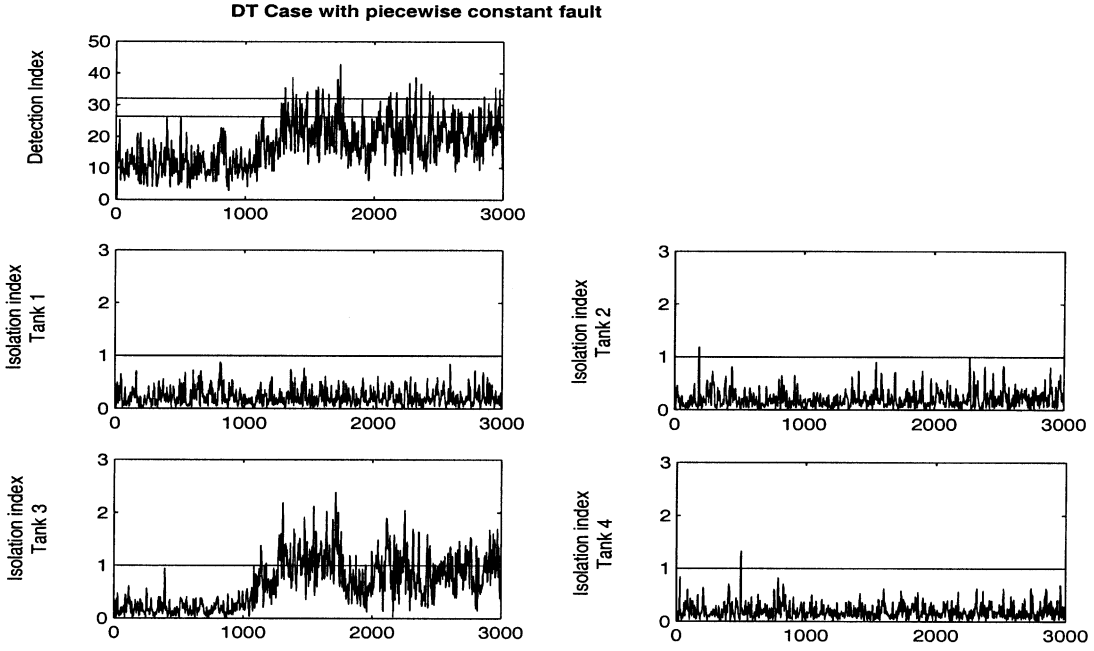

Fig. 9. The DT residual models can detect and isolate a leak with a ZOH in tank 3, with relatively poor performance.

systems has been proposed. When the number of outputs is not less than the system order, any single and multiple faults can be uniquely isolated by the use of n SRVs, where one SRV is designed to be affected only by a single fault but unaffected by the other faults. The proposed scheme is computationally robust and efficient, without the need to separately identify the system matrices {**A**, **B**, **C**, **D**}. Furthermore, the proposed approach can be extended to the identification of the system matrices {**A**, **B**, **C**, **D**} if they are needed. The newly proposed approach is applied to a simulated water tank system, where detection and isolation of leaks in a single tank one at a time and in multiple tanks simultaneously has been successfully conducted. In addition, to make a comparison, the results of detecting and isolating leaks in the tanks using the PRM and

SRMs in the DT domain are also provided. It is shown that the DT residual models-based approach can just barely detect and isolate only the piece-wise constant fault. In contrast, the CT residual models-based scheme works well for detecting and isolating single as well as multiple leaks from any tanks.

Although the proposed identification approach has demonstrated its effectiveness in PFDI, several issues such as the consistency analysis, fault detectability and isolability analysis and robustness with respect to process uncertainties, deserve further investigation. Work in this aspect is in progress.

Acknowledgements

Financial aid from the Natural Science and Engineering Research Council (NSERC), Matrikon Inc., and the Alberta Science Research Authority (ASRA) of Canada towards the NSERC, Matrikon, ASRA senior industrial research chair program at the University of Alberta is gratefully acknowledged.

Appendix A. Calculation of $\mathbf{W_0}$

To calculate the eigenvectors of $\mathbf{\Gamma}_s^{\perp}\mathbf{G}_s\mathbf{G}_s^T$, we express the problem as:

$$\mathbf{\Gamma}_s^{\perp}\mathbf{G}_s\mathbf{G}_s^T\boldsymbol{\nu}_i = \lambda\boldsymbol{\nu}_i,$$

where ν_i is the ith column of the matrix $\mathbf{W}_0^T$, and λ_i is the associated eigenvalue.

Introducing $\boldsymbol{\nu}_i = \mathbf{\Gamma}_s^{\perp}\boldsymbol{\omega}_i$ leads to

$$\mathbf{\Gamma}_s^{\perp}\mathbf{G}_s\mathbf{G}_s^T\mathbf{\Gamma}_s^{\perp}\boldsymbol{\omega}_i = \lambda_i\mathbf{\Gamma}_s^{\perp}\boldsymbol{\omega}_i.$$

Further, since $\mathbf{\Gamma}_s^{\perp}$ is idempotent, we obtain

$$\mathbf{\Gamma}_s^{\perp}\mathbf{G}_s\mathbf{G}_s^T\mathbf{\Gamma}_s^{\perp}\mathbf{\Gamma}_s^{\perp}\boldsymbol{\omega}_i = \lambda_i\mathbf{\Gamma}_s^{\perp}\boldsymbol{\omega}_i$$

and

$$\mathbf{\Gamma}_s^{\perp}\mathbf{G}_s\mathbf{G}_s^T\mathbf{\Gamma}_s^{\perp}\boldsymbol{\nu}_i = \lambda_i\boldsymbol{\nu}_i.$$

Hence, $\boldsymbol{\nu}_i$ is the ith eigenvector corresponding to the ith non-zero eigenvalue of $\mathbf{\Gamma}_s^{\perp}\mathbf{G}_s\mathbf{G}_s^T\mathbf{\Gamma}_s^{\perp}$. Eventually, $\mathbf{W}_0^T$ = Eigenvectors associated with all the non-zero eigenvalues of the matrix $\mathbf{\Gamma}_s^{\perp}\mathbf{G}_s\mathbf{G}_s^T\mathbf{\Gamma}_s^{\perp}$.

Appendix B. Consistent estimation of {A; B; C; D}

In many scenarios, e.g. in the design of controllers, knowledge of the system matrices is crucial. Therefore, we discuss the identification of $\{\mathbf{A}_T;\mathbf{B}_T;\mathbf{C}_T;\mathbf{D}_T\}$ from $\mathbf{\Gamma}_s$ and $\mathbf{W}_0\mathbf{H}_s$ in this Appendix. Note that the identified system matrices $\{\mathbf{A}_T, \mathbf{B}_T, \mathbf{C}_T, \mathbf{D}_T\}$ are in fact the similarity transformations of the original ones, e.g.

$$\mathbf{A}_T = \mathbf{T}\mathbf{A}\mathbf{T}^{-1}, \mathbf{B}_T = \mathbf{T}\mathbf{B}, \mathbf{C}_T = \mathbf{C}\mathbf{T}^{-1}, \mathbf{D}_T = \mathbf{D},$$

where $\mathbf{T} \in \Re^{n\times n}$ is a non-singular matrix.

The estimate of $\mathbf{C}_T$ is simply equal to the first m rows of $\mathbf{U}_l(:;\,1:n)$, i.e., $\mathbf{C}_T = \mathbf{U}_l(1:m;\,1:n)$. On the other hand, observing that

$$\mathbf{U}_l(m+1:m_s,1:n) = \begin{bmatrix}\mathbf{C}_T\mathbf{A}_T\\ \vdots\\ \mathbf{C}_T\mathbf{A}_T^s\end{bmatrix} = \begin{bmatrix}\mathbf{C}_T\\ \vdots\\ \mathbf{C}_T\mathbf{A}_T^{s-1}\end{bmatrix}\mathbf{A}_T = \mathbf{U}_l(1:m_s-m,1:n)\mathbf{A}_T$$

one arrives at

$$\mathbf{A}_T = \mathbf{U}_l^{+}(1:m_s-m,1:n)\mathbf{U}_l(m+1:m_s,1:n)$$

In order to estimate $\mathbf{B}_T$ and $\mathbf{D}_T$, we have to construct a series of matrix equations from $\mathbf{W}_0\mathbf{H}_s$. As illustrated in Section 2.2, in matrix $\mathbf{H}_s$ all the elements above the main diagonal are zero. As a consequence, for each column block with q columns in $\mathbf{W}_0\mathbf{H}_s$, it can be clearly seen, $\forall\, i=0$, that

$$\begin{aligned}&\mathbf{W}_0(:,mi+1:m_s)\mathbf{H}_s(:,mi+1:m_s,qi+1:qi+q)\\ &= \mathbf{W}_0(:,mi+1:m_s)\begin{bmatrix}\mathbf{I}_m & \mathbf{0}\\ \mathbf{0} & \mathbf{\Gamma}_{s-1-i}\mathbf{T}^{-1}\end{bmatrix}\begin{bmatrix}\mathbf{D}_T\\ \mathbf{B}_T\end{bmatrix}\\ &= \mathbf{W}_0(:,mi+1:m_s)\begin{bmatrix}\mathbf{I}_m & \mathbf{0}\\ \mathbf{0} & \mathbf{U}_l(1{:}m_s{-}mi{-}m,\,1{=}n)\end{bmatrix}\\ &\begin{bmatrix}\mathbf{D}_T\\ \mathbf{B}_T\end{bmatrix}\end{aligned} \tag{28}$$

where $\mathbf{\Gamma}_{-1} = 0$.

Denoting

$$\mathbf{N}_i = \mathbf{W}_0(:,mi+1:m_s)\mathbf{H}_s(mi+1:m_s,qi+1:qi+q)$$

and

$$\mathbf{M}_i = \mathbf{W}_0(:,mi+1:m_s) \times \begin{bmatrix}\mathbf{I}_m & 0\\ 0 & \mathbf{U}_l(1:m_s-mi-m,1=n)\end{bmatrix},$$

from Eq. (28) we establish the following set of equations:

$$\begin{bmatrix} \mathbf{N}_0 \\ \mathbf{N}_1 \\ \vdots \\ \mathbf{N}_s \end{bmatrix} = \begin{bmatrix} \mathbf{M}_0 \\ \mathbf{M}_1 \\ \vdots \\ \mathbf{M}_s \end{bmatrix} \begin{bmatrix} \mathbf{D}_T \\ \mathbf{B}_T \end{bmatrix} \tag{29}$$

Therefore, the least squares (LS) solution to $\{\mathbf{B}_\mathrm{T}\,;\,\mathbf{D}_\mathrm{T}\}$ is

$$\begin{bmatrix} \mathbf{D}_T \\ \mathbf{B}_T \end{bmatrix} = \left[\sum_{j=0}^{s} \mathbf{M}_j^T \mathbf{M}_j\right]^{-1} \left[\sum_{j=0}^{s} \mathbf{M}_j^T \mathbf{N}_j\right].$$

References

[1] D. Bauer, Order estimation for subspace methods, Automatica 37 (2001) 1561–1573.

[2] T. Chen, B. Francis, Optimal Sampled-Data Control Systems, Springer, Berlin, 1995.

[3] C. Chou, R. Johansson, M. Verhaegen, Continuous-time identification of SISO systems, IEEE Trans. Signal Processing 47 (1999) 349–362.

[4] C. Chou, M. Verhaegen, Subspace algorithms for the identification of multivariable dynamic errors-in-variables models, Automatica 33 (1997) 1857–1869.

[5] E. Chow, A. Willsky, Analytical redundancy and the design of robust failure detection systems, IEEE Trans. Auto. Cont 29 (1984) 603–614.

[6] P. Frank, Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy—a survey and some new results, Automatica 26 (1990) 459–474.

[7] W. Ge, C. Fang, Detection of faulty components via robust observation, Int. J. Control 47 (1988) 581–599.

[8] J. Gertler, D. Singer, Augmented models for statistical fault isolation in complex dynamic systems, in: Proc. of the American Control Conference, 1985, pp. 317–322.

[9] J. Gertler, D. Singer, A new structural framework for parity equation based failure detection and isolation, Automatica 26 (1990) 381–388.

[10] K. Johansson, The quadruple-tank process: a multivariable laboratory process with an adjustable zero, IEEE Trans. Cont. Sys. Tech. 8 (2000) 456–465.

[11] R. Johansson, M. Verhaegen, C. Chou, Stochastic theory of continuous-time state-space identification, IEEE Trans. Signal Processing 47 (1999) 41–51.

[12] W. Li, S. Shah, Structured residual vector-based approach to sensor fault detection and isolation, Journal of Process Control 12 (2002) 429–443.

[13] B. Liu, J. Si, Fault isolation filter design for linear time-invariant systems, IEEE Trans. Auto. Cont 21 (1997) 704–707.

[14] L. Ljung, System Identification: Theory for the User, Prentice-Hall, 1987.

[15] M. Massoumnia, A geometric approach to the synthesis of failure detection filter, IEEE Trans. Auto. Cont 31 (1986) 1867–1888.

[16] R. Middleton, G. Goodwin, Digital Control and Estimation, Prentice-Hall, 1990.

[17] M. Moonen, B. DeMoor, L. Vandenberghe, J. Vandewalle, On and off-line identification of linear state-space models, Int. J. of Control 49 (1989) 219–232.

[18] P. Overschee, B. De Moor, N4SID: subspace algorithms for the identification of combined deterministic-stochastic systems, Automatica 30 (1994) 75–93.

[19] P. Overschee, B. De Moor, A unifying theorem for three subspace system identification algorithms, Automatica 31 (1995) 1853–1864.

[20] P. Overschee, B. De Moor, Subspace Identification for Linear Systems-Theory, Implementation, Applications, Kluwer Academic Publisher, 1996.

[21] S. Qin, W. Li, Detection and identification of faulty sensors in dynamic processes, AIChE J. 47 (2001) 1581–1593.

[22] S. Sagara, Z. Zhao, Numerical integration approach to on-line identification of continuous-time systems, Automatica 26 (1990) 63–74.

[23] M. Verhaegen, P. Dewilde, Subspace model identification, part i: the output-error state-space model identification class of algorithms. Int. J. Control 56 (1992) 1187–1210.

[24] M. Verhaegen, P. Dewilde, Subspace model identification, part ii: analysis of the elementary output-error state-space model identification algorithm. Int. J. Control 56 (1992) 1211–1241.

[25] M. Viberg, Subspace-based methods for the identification of linear time-invariant systems, Automatica 31 (1995) 1835–1851.

[26] L. Wang, P. Gawthrop, On the estimation of continuous transfer functions, Int. J. Control 74 (2001) 889–904.

[27] J. White, J. Speyer, Detection filter design: spectral theory and algorithms, IEEE Trans. Auto. Cont 32 (1987) 593–603.

[28] A. Whitfield, N. Messali, Integral-equation approach to system identification, Int. J. Control 45 (1987) 1431–1445.

[29] A. Willsky, A survey of design methods for failure detection in dynamic systems, Automatica 12 (1976) 601–611.