MODEL BASED FAULT DETECTION FOR TWO-DIMENSIONAL SYSTEMS

by

Zhenheng Wang

Thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy (PhD) in Natural Resources Engineering

School of Graduate Studies
Laurentian University
Sudbury, Ontario, Canada

THESIS DEFENCE COMMITTEE/COMITÉ DE SOUTENANCE DE THÈSE
Laurentian Université/Université Laurentienne
School of Graduate Studies/École des études supérieures

Title of Thesis Titre de la thèse	MODEL BASED FAULT DETECTION FOR TWO-DIMENSIONAL SYSTEMS		
Name of Candidate Nom du candidat	Wang, Zhenheng		
Degree Diplôme	Doctor of Philosophy		
Department/Program Département/Programme	Natural Resources Engineering	Date of Defence Date de la soutenance	March 27, 2014

APPROVED/APPROUVÉ

Thesis Examiners/Examinateurs de thèse:

Dr. Helen Shang
(Supervisor/Directrice de thèse)

Dr. Junfeng Zhang
(Committee member/Membre du comité)

Dr. Brahim Chebbi
(Committee member/Membre du comité)

Dr. Anis Farah
(Committee member/Membre du comité)

Dr. Biao Huang
(External Examiner/Examinateur externe)

Dr. Kalpdrum Passi
(Internal Examiner/Examinateur interne)

Approved for the School of Graduate Studies
Approuvé pour l'École des études supérieures
Dr. David Lesbarrères
M. David Lesbarrères
Director, School of Graduate Studies
Directeur, École des études supérieures

ACCESSIBILITY CLAUSE AND PERMISSION TO USE

Abstract

Fault detection and isolation (FDI) are essential in ensuring safe and reliable operations in industrial systems. Extensive research has been carried out on FDI for one dimensional (1-D) systems, where variables vary only with time. The existing FDI strategies are mainly focussed on 1-D systems and can generally be classified as model based and process history data based methods. In many industrial systems, the state variables change with space and time (e.g., sheet forming, fixed bed reactors, and furnaces). These systems are termed as distributed parameter systems (DPS) or two dimensional (2-D) systems. 2-D systems have been commonly represented by the Roesser Model and the F-M model. Fault detection and isolation for 2-D systems represent a great challenge in both theoretical development and applications and only limited research results are available.

In this thesis, model based fault detection strategies for 2-D systems have been investigated based on the F-M and the Roesser models. A dead-beat observer based fault detection has been available for the F-M model. In this work, an observer based fault detection strategy is investigated for systems modelled by the Roesser model. Using the 2-D polynomial matrix technique, a dead-beat observer is developed and the state estimate from the observer is then input to a residual generator to monitor occurrence of faults. An enhanced realization technique is combined to achieve efficient fault detection with reduced computations. Simulation results indicate that the proposed method is effective in detecting faults for systems without disturbances as well as those affected by unknown disturbances.

The dead-beat observer based fault detection has been shown to be effective for 2-D systems but strict conditions are required in order for an observer and a residual generator to exist. These strict conditions may not be satisfied for some systems. The effect of process noises are also not considered in the observer based fault detection approaches for 2-D systems. To overcome the disadvantages, 2-D Kalman filter based fault detection algorithms are proposed in the thesis.

A recursive 2-D Kalman filter is applied to obtain state estimate minimizing the estimation error variances. Based on the state estimate from the Kalman filter, a residual is generated reflecting fault information. A model is formulated for the relation of the residual with faults over a moving evaluation window. Simulations are performed on two F-M models and results indicate that faults can be detected effectively and efficiently using the Kalman filter based fault detection.

In the observer based and Kalman filter based fault detection approaches, the residual signals are used to determine whether a fault occurs. For systems with complicated fault information and/or noises, it is necessary to evaluate the residual signals using statistical techniques. Fault detection of 2-D systems is proposed with the residuals evaluated using dynamic principal component analysis (DPCA). Based on historical data, the reference residuals are first generated using either the observer or the Kalman filter based approach. Based on the residual time-lagged data matrices for the reference data, the principal components are calculated and the threshold value obtained. In online applications, the T^2 value of the residual signals are compared with the threshold value to determine fault occurrence. Simulation results show that applying DPCA to evaluation of 2-D residuals is effective.

Key words: Two dimensional systems, Fault detection, Kalman filter, Observer, Polynomial theory

Acknowledgements

I would like to express my sincere gratitude to my supervisor Dr. Helen Shang for her patient guidance, continuous encouragement and unconditional support both mentally and financially throughout the four years of PhD study. Most importantly, she gave me the opportunity to do my PhD research under her supervision.

I would like to thank my advisory committee, Dr. Brahim Chebbi, Dr. Junfeng Zhang and Dr. Anis Farah, for their time in helping me to finish my Ph.D work, and reading and reviewing the thesis to improve and enhance its quality. I am also grateful for the internal examiner Dr. Kalpdrum Passi from Department of Mathematics & Computer Science for spending his time in reviewing the thesis and providing valuable suggestions. Insightful suggestions from the external examiner Dr. Biao Huang from University of Alberta are especially appreciated.

The Natural Sciences and Engineering Research Council of Canada (NSERC) is gratefully acknowledged for their financial support.

Contents

List of Figures

List of Abbreviations

1-D	One-Dimensional
2-D	Two-Dimensional
AEM	Abnormal Event Management
ANN	Artificial Neural Network
ARMAX	Autoregressive-moving-average Exogenous Model
ARX	Autoregressive Exogenous Model
DPCA	Dynamic Principal Component Analysis
DPS	Distribute Parameter Systems
DTW	Dynamic Time Warping
EKF	Extend Kalman Filter
F-M	Fornasini & Maichesini
FDI	Fault Detection and Isolation
FLC	Factor Left Coprime
FRC	Factor Right Coprime
GLCD	Greatest Left Common Divisor
GLR	Generalized Likelihood Ratio
GRCD	Greatest Right Common Divisor
KF	Kalman Filter
LR	Likelihood Ratio
LTI	Linear Time Invariant
MFD	Matrix Fraction Description
MLA	Minimal Left Annihilator
MLC	Minor Left Coprime
MPCA	Multiway Principal Component Analysis

MRC	Minor Right Coprime
PBH	Observability matrix of process
PCA	Principal Component Analysis
PDE	Partial Derivative Equation
rMP	right Minor Prime
rVP	right Variety Prime
rZP	right Zero Prime
UIO	Unknown Input Observer
ZLC	Zero Left Coprime
ZRC	Zero Right Coprime
ZRP	Zero Right Prime

Chapter 1

Introduction

1.1 Fault detection and model based fault detection

The term fault is generally defined as a departure from an acceptable range of an observed variable or a calculated parameter associated with a process (Himmelblau, 1978). Fault detection and diagnosis deals with the timely detection, diagnosis and correction of abnormal conditions of faults in a process. Early detection and diagnosis of process faults while the plant is still operating in a controllable region can help avoid abnormal event progression and reduce productivity loss. Since the petrochemical industries lose an estimated 20 billion dollars every year, they have rated Abnormal Event Management (AEM), of which fault detection and diagnosis is the central component, as their number one problem that needs to be solved (Nimmo, 1995). Hence, there is considerable interest in this field now from industrial practitioners as well as academic researchers.

1.1.1 Existing research and techniques

Fault detection and isolation (FDI) is a relatively young research field since 90s' but has become an integral component of system engineering. The overall concept of FDI consists of the following two components:

- Fault detection: detection of the occurrence of faults in the functional units of process, which could lead to undesirable or intolerable compromise of a whole plant.

- Fault isolation: localization or classification of different faults based on historical data.

Extensive research has been carried out on FDI of one dimensional (1-D) systems. There is an abundance of literature on process fault detection ranging from analytical methods to artificial intelligence and statistical approaches (Wu et al., 2012). A classification of FDI techniques is given in Figure 1.1. The FDI methods can also be classified as model based methods and process history data based methods. Using the process history data, multivariate statistical process modeling methods (see e.g., Jackson (1991)) have been extensively used in process monitoring and fault detection (Ku et al., 1995; Lee et al., 2004; Kourti and MacGregor, 1996). The methods are data-driven and one of the most commonly used tools is principal component analysis (PCA) (Wise and Gallagher, 1996; Russell et al., 2000). Multiway PCA (MPCA) and dynamic PCA (DPCA) are extensions of PCA for batch processes (Nomikos and MacGregor, 1994; Wise et al., 1999), and has been used in various applications (Gameroa et al., 2006; Khosravi et al., 2009). These methods are efficient in detecting abnormal events, but use of the contribution plots in fault diagnosis may sometimes lead to inaccurate results. Fault diagnosis can also be conducted by comparing the real time signals with a database of reference signals representing various faults. The challenge is that two similar signals are often slightly different and do not match each other perfectly. Dynamic time warping (DTW) has been used to overcome the challenge and it was used in fault detection and diagnosis of batch processes or transition periods (Kassidas et al., 1998; Srinivasan and Qian, 2005, 2006; Dai and Zhao, 2011). A hybrid FDI method using online DTW in combination of PCA and SymCure reasoning under the G2 Optegrity was recently developed for chemical process startups (Wang et al., 2012). Artificial neural networks (ANNs) have also been applied in fault detection and diagnosis with the available online training methods (Ruiz et al., 2000; Zhang, 2006). This approach requires a large amount of training samples, but signal samples with faults are usually very limited in the industrial applications. Other techniques such as petri net and expert systems have been proposed to solve the problem for batch processes (Power and Bahri, 2004; Mehranbod et al., 2005).

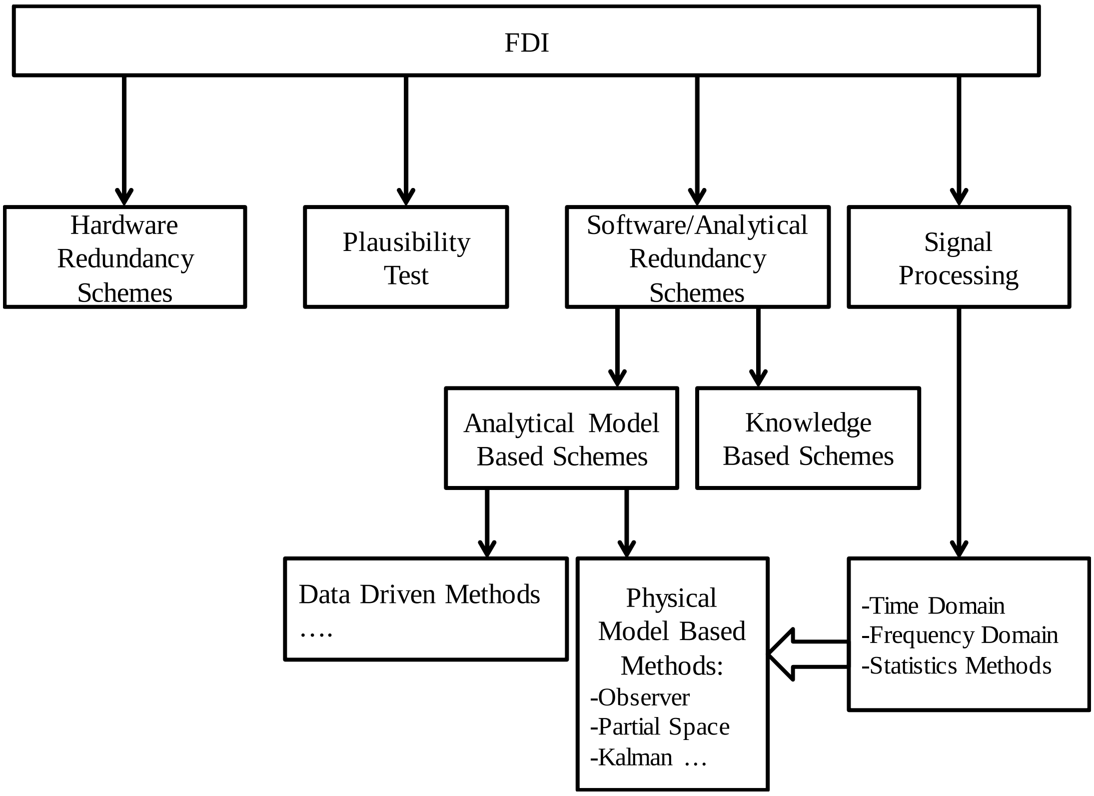


Figure 1.1: The classification of FDI methods

Analytical model fault diagnosis strategies have attracted interests of researchers (Frank, 1996). A priori process knowledge plays a crucial role in diagnostic decision-making and could be in the form of material and energy balances (Mehra and Peschon, 1971; Kramer, 1987), suitably transformed process models (Willsky and Jones, 1976; Gertler, 1991; Frank, 1990) or distribution information in the measurement space (Hoskins and Himmelblau, 1988). The first-principles models (also classified as macroscopic transport phenomena model (Himmelblau, 1978)) have not been very popular in fault diagnosis studies because of the computational complexity in utilizing these models in real-time fault diagnostic systems and the difficulty in developing these models. The models that have been heavily investigated in fault diagnosis are input-output or state space models. Faults in the state space framework are described in a series of papers by Gertler (1991, 1993) and in the recent books on fault detection and diagnosis written by (Gertler, 1998; Chen and Patton, 1999; Russell et al., 2000).

Model-based FDI requires a process model running in parallel to the process. The analytical schemes for fault diagnosis are basically signal processing techniques using state estimation, parameter estimation, adaptive filtering and so on (Venkatasubramanian et al., 2003). The key component in these schemes is diagnostic observers for residue generation. Generation of diagnostic observers for nonlinear systems has also been considered to a certain extent in the literature (Clark, 1979; Frank, 1990; Massoumnia, 1996). Most of the work on observer design focuses on the generation of residuals for dynamic systems with satisfactory decoupling properties. Another solution (Basseville, 1988; Willsky and Jones, 1976) to the fault diagnosis problem in such systems entails monitoring the innovation process or the prediction errors. The objective is to design a state estimator with minimum estimation error. It involves the use of optimal state estimate, e.g. the Kalman filter, which is designed on the basis of the system model in its normal operating mode. The statistical analysis of Kalman filter was pioneered by Willsky and Jones (1976) and further explored by Basseville (1993) and the references therein. It has been shown that a bank of Kalman filters (Basseville, 1993) designed on the basis of all the available possible system models under all possible changes can be used for the isolation purpose. Literature (Fathi et al., 1993) included adaptive analytical redundancy models in the diagnostic reasoning loop of knowledge-based systems. The modified extended Kalman filter (EKF) is used in designing local detection filters in their work. In a recent work, literature (Chang and Hwang, 1998) explored the possibility of using suboptimal EKF in order to enhance computation efficiency without sacrificing diagnosis accuracy.

The analytical model based FDI is illustrated in Figure 1.2. In this scheme, a process model is constructed according to the process input and output, the certain features are extracted, which could be used as residual generator. By comparing with the normal behavior, the detection results are obtained. The residual signal is generally corrupted with disturbances and uncertainties caused by parameter changes. To achieve a successful fault detection based on the available residual signal, a widely accepted way is to generate such a feature of the residual signal, by which we are able to distinguish the faults from the disturbances and uncertainties. Residual evaluation and threshold setting serve for this purpose. A decision on the possible occurrence of a fault is then made by means of a simple comparison between the residual feature and the threshold. The fault isolation may require some analytical schemes to evaluate the historical faulty data in order to locate the fault type and associated information.

1.1.2 Residual generation

In model based FDI, residual generation is to compare system measurements with the model based reconstruction of system outputs. Design of a residual generator requires that a residual signal is an indicator of possible faults. Process faults usually cause changes in the state variables and/or the model parameters. Based on a process model, one can estimate the unmeasurable state variables using state estimation and parameter estimation methods. State estimation is often required in residual generation. Kalman filters or observers have been widely used for state estimation. Techniques relying on parity equations for residual generation have also been developed. In this subsection, the basic idea of model based residual generation is illustrated using a fault detection filter, which is the first type of observer based residual generators proposed

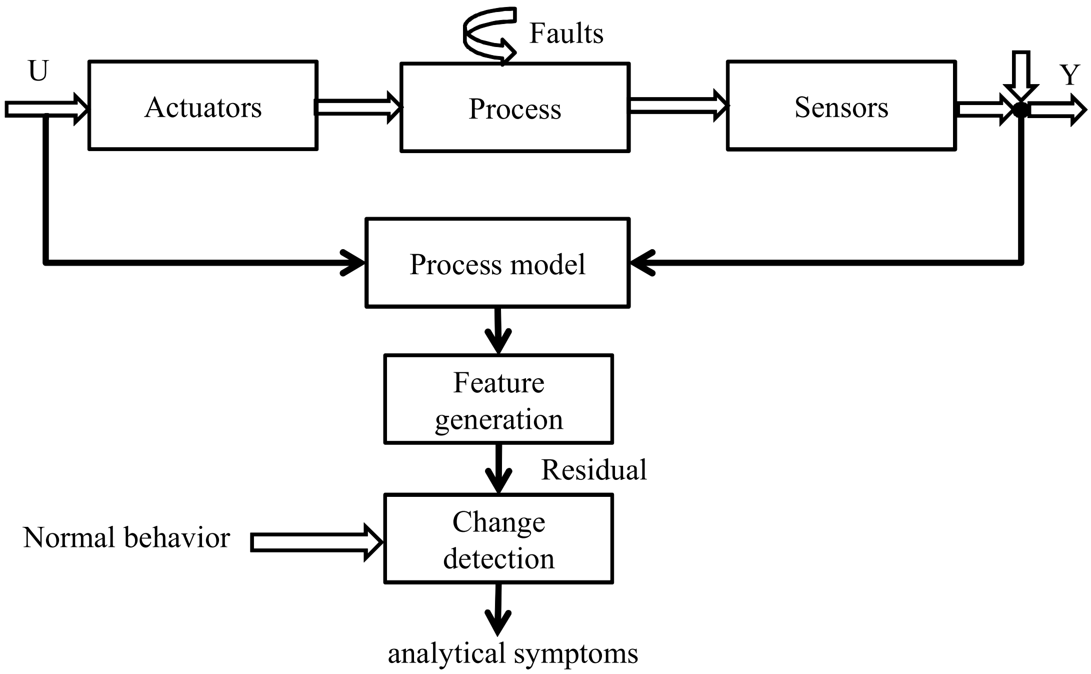


Figure 1.2: The diagram of a model-based FDI

by Beard and Jones in the early 1970s.

Consider a system described by the following discrete time state-space equations:

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) + Ff(t) \\ y(t) &= Cx(t) \end{aligned}$$

where t indicates the discrete time, x, u, y and f indicate the state, input, output and fault, respectively. A, B, C and F are given system matrices of proper dimension. A state observer can be designed:

$$\hat{x}(t+1) = A\hat{x}(t) + Bu(t) + L[y(t) - C\hat{x}(t)]$$

where $\hat{x}$ denotes the estimate of the state x. A residual signal can be simply defined as:

$$r(t) = y(t) - \hat{y} = y(t) - C\hat{x}(t)$$

Denoting state estimation error as $e(t) = x(t) - \hat{x}(t)$, $e(t)$ and $r(t)$ can then be expressed as:

$$\begin{aligned} e(t+1) &= (A - LC)e(t) + Ff(t) \\ r(t) &= Ce(t) \end{aligned}$$

If the absolute values of the eigenvalues of $(A - LC)$ are less than 1, $e(t) \rightarrow 0$ as $t \rightarrow \infty$. As a result, in fault-free case, i.e., $f = 0$, the estimation error and consequently the residual track the process. When a fault f occurs, the error and the residual carry the information of the faults and faults are reflected in estimation error and residual.

Since the early development of the fault detection filter, a variety of different observer based approaches have been proposed such as those that can deal with unknown disturbance decoupling. In the observer based residual generation, observers are designed such that they are sensitive to a subset of faults while insensitive to the remaining faults and the unknown inputs. In a fault-free case, the observers track the process closely with small residuals. If a fault occurs, the observers that are sensitive to the fault deviate from the process significantly and result in residuals of large magnitude.

Residual generation is an important step for model based fault detection. With the generated residual signal, residual evaluation and threshold setting is then performed to detect and diagnose the faults.

1.1.3 Residual evaluation and threshold setting

The residual signal is generally corrupted with disturbances and uncertainties caused by parameter changes. To achieve a successful fault detection based on the available residual signal, further efforts are needed. A widely accepted way is to generate such a feature of the residual signal, by which we are able to distinguish the faults from the disturbances and uncertainties. Residual evaluation and threshold setting serve for this purpose. A decision on the possible occurrence of a fault is then made by means of a simple comparison between the residual feature and the threshold.

Two strategies including the statistical testing and the norm based testing are often used for residual evaluation. In the norm based residual evaluation, an evaluation function that gives the mathematical feature of the residual signal is first defined and the decision on fault detection is made based on comparison of the evaluation function with a threshold value. For a given residual sequence, the peak value is one of the most straightforward evaluation function. The peak value of a discrete time residual signal $r(k)$ is defined as:

$$J_{\text{peak}} = \|r\|_{\text{peak}} := \sup_{k \geqslant 0} \|r(k)\|$$

with the norm of $r(k)$ defined by:

$$\|r(k)\| = (\sum_{i=1}^{k_{max}} r_i^2(k))^{1/2}$$

where k_{max} is the length of the signal sequence. Based on the peak value of r, faults can be detected by comparing it with a threshold value:

$$J_{\text{peak}} > J_{\text{th,peak}} \Rightarrow \text{alarm, fault is detected}$$

$$J_{\text{peak}} < J_{\text{th,peak}} \Rightarrow \text{no alarm, fault-free}$$

where $J_{\text{th,peak}}$ is the threshold for J_{peak}. Determination of a threshold is to find out the tolerant limit for disturbances and model uncertainties under fault-free operation conditions. $J_{\text{th,peak}}$ in the above equation is defined by

$$J_{\text{th,peak}} = \sup_{\text{fault free}} J_{\text{peak}}$$

Different evaluation functions, such as the trend value and average value, are also available for residual evaluation. In practice, the limit monitoring and trend analysis have been very popular in fault detection due to their ease of applications.

In addition to the norm based residual evaluation, statistical techniques have been successfully applied for residual evaluation and threshold setting. Statistical distribution functions, e.g., t , χ^2 and F distributions, have been adopted to calculate the control limit. Likelihood ratio (LR) and generalized likelihood ratio (GLR) methods are one of the most popular statistical methods for fault detection. The GLR method is applied to evaluate the residual and calculate the threshold for the Kalman filter based fault detection in this thesis. A good review on the statistical approach based residual evaluation can be found in the book by (Ding, 2008).

1.2 Two dimensional systems

In many industrial systems, the state variables change with space and time (e.g., sheet forming, fixed bed reactors, and furnaces). These systems are termed as distributed parameter systems (DPS) or two dimensional (2-D) systems. Motivated by interest in processing two-dimensional data, such as sampled seismic data, gravitational and magnetic maps, and 2-D images (Fornasini and Maichesini, 1977), 2-D state space models were proposed as a means to unify the study of these linear systems. The Fornasini & Marchesini (F-M) (Fornasini and Maichesini, 1977) and the Roesser models (Roesser, 1975) were presented in the 1970's and since then they have commonly been adopted to describe discrete-time, discrete-space linear 2-D systems, and research on 2-D systems has been carried out on the topics commonly investigated in one-dimensional (1-D) settings. The concept of 2-D transfer function, controllability and observability were defined for the F-M model and Roesser model (Rikus, 1979; Criftcibasi and Yuksel, 1983). The model representations such as ARX and ARMAX in 1-D notation was extended to 2-D cases (Isaksson, 1993). Identification techniques for 2-D systems were investigated using least square and its extensions (Zhao and Yu, 1993) and subspace identification (Ramos, 1993; Ramos et al., 2011). Active research has also been carried out on filters (Woods and Radewan, 1977; Makoto and Sumihisa, 1991; Wang, 1988; Wu et al., 2008; Li and Gao, 2012; Xu and Zou, 2011; Li and Gao, 2013), state estimation (Kaczorek, 2001; Ntogramatzidis and Cantoni, 2012; Xu et al., 2012), realization (Fornasini and Maichesini, 1988; Antoniou et al., 1988; Birgit, 2002) and control design (Yang et al., 2006; Xu and Zou, 2011; Ye et al., 2011; Li et al., 2012) of 2-D systems.

1.2.1 2-D state space models

The Roesser model is defined by (Roesser, 1975):

$$\begin{aligned}
\begin{bmatrix} x^h(i+1,j) \\ x^v(i,j+1) \end{bmatrix} &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x^h(i,j) \\ x^v(i,j) \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u(i,j) \\
y(i,j) &= \begin{bmatrix} C_1 & C_2 \end{bmatrix} \begin{bmatrix} x^h(i,j) \\ x^v(i,j) \end{bmatrix} + Ku(i,j)
\end{aligned} \tag{1.1}$$

where i and j indicate the horizontal (or spacial) and vertical (or time) indices, x^h is the state variable evolving horizontally (or spatially), x^v is the state variable evolving vertically (or in time), u and y indicate the process input and output, respectively, A's, B's, C's and K are the real matrices of suitable dimensions. The system can be described in abbreviation by $\{A_{11}, A_{12}, A_{21}, A_{22}, B_1, B_2, C_1, C_2, K\}$.

The F-M model is defined by (Fornasini and Maichesini, 1977):

$$\begin{aligned}
x(i+1,j+1) =& A_1 x(i+1,j) + A_2 x(i,j+1) \\
& + B_1 u(i+1,j) + B_2 u(i,j+1) \\
y(i,j) =& C(i,j)x(i,j) + K_u u(i,j)
\end{aligned} \tag{1.2}$$

Another form of F-M model (Attasi, 1976) is described by:

$$\begin{aligned}
x(i+1,j+1) =& A_1 x(i+1,j) + A_2 x(i,j+1) + A_3 x(i,j) \\
& + Bu(i,j) \\
y(i,j) =& Cx(i,j) + K_u u(i,j)
\end{aligned} \tag{1.3}$$

where x, u and y indicate the process state, input and output, respectively, A_k's, B_k's($k = 1, 2$), C's and K_u are the real matrices of suitable dimensions.

A practical example of Roesser model and F-M model can be found in a forced-flow steamjacketed tabular heat exchanger described by a linear hyperbolic PDE written as:

$$\frac{\partial(T,t)}{\partial t} + a\frac{\partial T(,t)}{\partial x} = bT(x,t) + cu(x,t), \tag{1.4}$$

where T denotes the temperature of the heat exchanger within the tube, the input u could be the jacket temperature. Its analytical solution, found using the method of characteristics, is:

$$T(x_o + a\Delta t, t_o + \Delta t) = T(x_o, t_o)e^{b\Delta t} + (\frac{c}{b}e^{b\Delta t} - \frac{c}{b})u(x_o, t_o). \tag{1.5}$$

By setting $\Delta x = a\Delta t$, Equation (1.5)is in the form of an F-M model:

$$T(i+1, j+1) = e^{b\Delta t}T(i,j) + (\frac{c}{b}e^{b\Delta t} - \frac{c}{b})u(i,j). \tag{1.6}$$

Define $x^h(i,j) = T(i,j)$ and $x^v(i,j) = T(i+1,j)$. Equation (1.5) can also be expressed as the Roesser model form

$$\begin{aligned} \begin{bmatrix} x^h(i+1,j) \\ x^v(i,j+1) \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ e^{b\Delta t} & 0 \end{bmatrix} \begin{bmatrix} x^h(i,j) \\ x^v(i,j) \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{c}{b}e^{b\Delta t} - \frac{c}{b} \end{bmatrix} u(i,j) \\ y(i,j) &= \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x^h(i,j) \\ x^v(i,j) \end{bmatrix} + Du(i,j) \end{aligned} \tag{1.7}$$

The Roesser Model describes the time domain as the horizontal vector and the spatial domain as the vertical vector, and a process is described by a series of two dimensional coordinates. The F-M model has two dimensional coordinates but does not have two directional vectors. Its recursive forward process is determined by several previous 2-D coordinates. So far, Roesser Model has been used in the areas of control and realization while F-M model used in estimation and filter problems. Fornasini and Marchesini provided the algebraic view of a 2-D system. In their interpretation, the state space arises from the factorization of the 2-D input-output map. They indicated that the main difference between 1-D and 2-D systems lies in that we can introduce a global state and a local state in 2-D cases. For the Roesser model, on the other hand, the "circuit approach" has been used for the state space model's realization (Criftcibasi and Yuksel, 1983). Although the two dimensional systems are widely available in industrial processes, the control theory and methods in 2-D form, especially in fault detection and isolation domain, represent a great challenge in both research and applications.

Both Roesser model and F-M model can be used to represent linear discrete 2-D systems. (Kung et al., 1977) showed that Roesser model was a more general model form for 2-D systems and F-M model represents a special case of the systems described by Roesser model. For a more generalized F-M model (Attasi, 1976) described by:

$$\begin{aligned} x(i+1,j+1) =& A_1x(i+1,j) + A_2x(i,j+1) + A_3x(i,j) \\ &+ Bu(i,j) \\ y(i,j) =& Cx(i,j) + K_uu(i,j) \end{aligned} \tag{1.8}$$

Let $\xi(i,j) = x(i,j+1) - A_1x(i,j)$ such that

$$\xi(i+1,j) = A_2x(i,j+1) + A_3x(i,j) + Bu(i,j) \tag{1.9}$$

Equation (1.8) can thus be expressed in the Roesser model form:

$$\begin{aligned} \begin{bmatrix} \xi(i+1,j) \\ x(i,j+1) \end{bmatrix} &= \begin{bmatrix} A_2 & A_3 + A_2A_1 \\ I_n & A_1 \end{bmatrix} \begin{bmatrix} \xi(i,j) \\ x(i,j) \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u(i,j) \\ y(i,j) &= \begin{bmatrix} 0 & C \end{bmatrix} \begin{bmatrix} \xi(i,j) \\ x(i,j) \end{bmatrix} \end{aligned} \tag{1.10}$$

Equation (1.2) is a special case when $A_3 = 0$.

On the other hand, the Roesser model can also be described in the F-M model form. By defining $x(i,j) = \begin{bmatrix} x^h(i+1,j) \\ x^v(i,j+1) \end{bmatrix}$, Equation (1.1) can then be written as:

$$\begin{aligned} x(i+1,j+1) =& \begin{bmatrix} 0 & 0 \\ A_{21} & A_{22} \end{bmatrix} x(i+1,j) + \begin{bmatrix} A_{11} & A_{12} \\ 0 & 0 \end{bmatrix} x(i,j+1) \\ &+ \begin{bmatrix} 0 \\ B_2 \end{bmatrix} u(i+1,j) + \begin{bmatrix} B_1 \\ 0 \end{bmatrix} u(i,j+1) \\ y(i,j) =& Cx(i,j) + Du(i,j) \end{aligned} \tag{1.11}$$

Despite their exchangeability, research has been focused on both Roesser and F-M models instead of replacing one by the other. In this thesis, fault detection methods for these two model forms are investigated using the 2-D observer based and Kalman filter based approaches.

In 1-D applications, research has been understandably focussed on the systems that are causal. This is reasonable as any current input can only affect the outputs in the future. In 2-D applications, the systems may not be causal in space. Due to complexity of 2-D systems, causality has, however, been assumed in both time and spatial directions as in Roesser and F-M models. In this thesis, it is assumed that the 2-D systems are causal in both time and spatial directions, and can be expressed by Roesser or F-M models.

1.2.2 2-D observers for state estimation and fault detection

Observers have been used for state estimate in the process control area for decades. Two kinds of observers are available: state observer and output observer. State observer adjust the state variables according to initial conditions and to the evolvement of the measured input and output signals. The observer-based estimation technique has been active in the area of control theory and engineering. System observability is an important prerequisite for the design of a state observer. In the early development stage of the observer-based estimate technique, system observability was considered as a necessary structural condition for the observer construction.

The 2-D dead-beat observer has been introduced into process estimate for a long time. But application of the dead-beat observer into fault detection and isolation in 2-D systems appeared only in the recent years, when literature (Bisiaco and Valcher, 2004, 2006; Bisiacco and Valcher, 2008) applied the dead-beat observer in fault detection and isolation (FDI) by constructing a residual generator. 2-D polynomial matrices are essential to construct the residual generator. The observer-based FDI is an approach of fault detection that requires the properties related with the two dimensional polynomial matrix theories. In the papers presented by (Bisiacco and Valcher, 2008), the unaffected FDI and the UIO (unknown input observer) FDI represent two different scenarios occurring in faulty 2-D F-M systems.

1.2.3 2-D Kalman filters

Filters have been widely applied in state estimation in many industrial fields. 1-D Kalman Filters, known for its high efficiency, has been successfully used in industrial processes. For 2-D systems, recursive filters have been adopted for state estimation. Some early developments such as that in (Celebi and Kurz, 1991; Hinamoto et al., 1997) were focused on the structural filtering for 2-D state space models. Although it is efficient and mathematically accessible, advantages of filtering have not fully realized.

Motivated by applications of 2-D state space models in 2-D signal and image processing, it is promising to extend the Kalman filter to 2-D cases. In a 2-D case, the enormous quantity of the data calls particularly for an efficient recursion processor. Originally, 2-D Kalman filters were introduced to handle image restoration and some signal processing related fields (Azimi-Sadjadi and Wong, 1987; Azimi-Sadjadi and Bannour, 1991; Suresh and Shenoi, 1981; Zhang and Steenaart, 1990; Kaufman et al., 1983). The concept of 2-D Kalman filter is fully examined by Woods (Woods and Radewan, 1977; Woods, 1979; Woods and Ingle, 1981). Unfortunately, the efforts to achieve a true recursive 2-D Kalman filter were of only limited success because of both the difficulty in establishing a suitable 2-D recursion model and the high dimension of the resulting state vector. In fact, a straightforward extension of one-dimensional Kalman filter techniques would result in a number of state variables proportional to N for the filtering of an $N \times N$ digital image (Sheng and Zou, 2007; Kwan and Lewis, 1999; Katayama and Kosaka, 1979). From 1970s, there have been many researches that are involved with attempting to introduce the 2-D Kalman filters, such as (Shanks et al., 1972; Woods and Radewan, 1977; Katayama and Kosaka, 1979; Sebek, 1992). However, to achieve a full recursive Kalman filter is still of limited success.

A new 2-D recursive Kalman filter structure was recently developed in (Sheng and Zou, 2007). In this thesis, a fault detection method based on the recursive Kalman filter is investigated.

1.3 Thesis organization and contributions

The thesis is divided into 6 chapters. In Chapter 1, an introduction and a literature review on FDI and two dimensional systems are presented. A background of polynomial theory is provided in Chapter 2. The contributions of this dissertation can be found in Chapters 3 through 5. Finally, the conclusions and future work are given in Chapter 6.

The contributions of this thesis are summarized as follows:

- **Chapter 3**: This chapter presents a development of dead beat observer based fault detection for 2-D systems described by the Roesser model. Based on a PBH observability matrix, a state observer is calculated and the obtained state estimate enters a residual generator to provide a residual signal reflecting the fault information. An efficient realization technique is incorporated to obtain the state space models of the observer and the residual generator. The proposed fault detection strategy is effective for systems without unknown disturbances as well as those with unknown disturbances.

- **Chapter 4**: This chapter presents a fault detection approach for systems described by the F-M model with noises. Using a 2-D recursive Kalman filter, a state estimate is obtained by minimizing the estimation variances and a residual signal is then obtained from an innovation process. The model for the residual over the evaluation window is formulated relating the residual with the faults. Evaluation of the residual signal is explored using the generalized likelihood ratio (GLR) test.

- **Chapter 5**: In Chapters 3 and 4, residual signals reflecting fault occurrences are obtained using an observer based residual generator or a Kalman filter based residual generator. To determine occurrences of faults more accurately, it is necessary to perform residual evaluation and threshold setting in 2-D systems. In Chapter 5, the fault detection method is proposed with the residual signals evaluated using dynamic principal component analysis. A historical reference data are used to obtain the DPCA models for the residuals. By calculating the T^2 value of the residual signals and comparing it with its threshold value, a fault is detected.

Chapter 2

Mathematical Background on Polynomial Matrices

In developing model based fault detection for 2-D systems, 2-D polynomial matrices have been used, especially for the observer based fault detection method. In this chapter, mathematics background on polynomial matrices in 2-D systems is introduced for the convenience of reading the subsequent chapters.

2.1 Polynomial matrices in 1-D systems

Since the sixties, polynomial matrices have constituted a fundamental tool for investigating the dynamics of a linear system and for designing feedback control laws (Rosenbrock, 1970). The input-output behavior of a system can be described as a polynomial matrix commonly called a transfer function. For a linear time invariant(LTI) system, the polynomial representation is a standard mathematical form in the frequency domain. Denoting $G_{yu}(p) \in \mathbb{R}$ is the transfer function of input vector $u \in \mathbb{R}$ and output vector $y \in \mathbb{R}$, the system can be described by:

$$y(p) = G_{yu}(p)u(p)$$

Among the different polynomial representations, Laurent polynomial is the most used one, especially in the process control domain.A Laurent polynomial with coefficients in a field $\mathbb{F}$ is defined as an expression of the form $p = \sum\limits_k p_k X^k, p_k \in \mathbb{F}$, where X is a formal variable, the summation index k is a integer (not necessarily positive) and only finitely many coefficients p_k are non-zero. Two Laurent polynomials (L-polynomial) are equal if their coefficients are equal. Such expressions can be added, multiplied, and brought back to the same form by reducing similar terms. Formulas for addition and multiplication are exactly the same as for the ordinary polynomials, with the only difference that both positive and negative powers of X can be present:

$$(\sum_i a_i X^i) + (\sum_i b_i X^i) = \sum_i (a_i + b_i) X^i$$

and

$$(\sum_i a_i X^i) \cdot (\sum_i b_i X^i) = \sum_k (\sum_{i,j;i+j=k} a_i b_j) X^k$$

Since only finitely many coefficients a_i and b_j are non-zero, all sums in effect have only finitely many terms, and hence represent Laurent polynomials. In fact, in LTI systems, most transfer functions can be expressed as L-polynomials. L-polynomials are widely used in behavior approaches of LTI systems.

Consider a discrete time state space model given by:

$$\begin{aligned} x(k+1) =& Ax(k) + Bu(k), x(0) = x_0 \\ y(k) =& Cx(k) + Du(k) \end{aligned}$$

The above state space model can be derived either by direct modeling or from a transfer function $G_{yu}(p) \in \mathbb{F}$. The latter process is called realization of the transfer function $G_{yu}(p) = C(pI - A)^{-1}B + D$ and denote by:

$$G_{yu}(p) = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

It is desirable to make (A, B, C, D) the minimal realization of $G_{yu}(p)$.

System analysis has been investigated using Laurent polynomial matrices. Many applications based on Laurent polynomial matrices are in the areas of realization, factorization, decoding and residual generators. In a 1-D context, the polynomial matrix algebra applied for solving the aforementioned problems is rather simple. Based on elementary transformations, efficient algorithms allow for a complete analysis of the system dynamics.

2.2 Polynomial matrices in 2-D systems

The use of polynomial matrices in 2-D system analysis and control began in the late seventies (Bisiacco, 1985; Morf et al, 1977; Guiver and Bose, 1985), while more recently Rocha and Willems (1991) resorted to polynomial matrices in two variables for introducing 2-D behaviors. The structure of a 2-D trajectories on $\mathbb{Z} \times \mathbb{Z}$ is endowed with the higher complexity exhibited by 2-D polynomial rings and matrices.

Let $\mathbb{F}$ be a field and denote by z the n-tuple $(z_1, z_2, \cdots, z_n)$and by z_i^c the $(n-1)$-tuple $(z_1, z_2, \cdots, z_{i-1}, z_{i+1}, \cdots, z_n)$, so that $\mathbb{F}[z, z^{-1}]$ and $\mathbb{F}[z_i^c, (z_i^{-1})^c]$ are shorthand notations for the L-polynomial rings in the indeterminates $z_1, \cdots, z_n$ and $z_1, \cdots, z_{i-1}, z_{i+1}, \cdots, z_n$, respectively, and $\mathbb{F}(z)$ denotes the field of rational functions with coefficients in $\mathbb{F}$. For the 2-D system, a polynomial ring is denoted as a ring which belong to $\mathbb{F}(z_1, z_2)$.

2.2.1 Primeness

A matrix $G \in \mathbb{F}[z, z^{-1}]^{p \times m}$ has rank r if it has a nonzero r-th order minor, whereas all its higher orders minors are zero. The rank of a matrix coincides with the dimensions of the $\mathbb{F}(z)$-spaces generated either by its rows or by its columns. By referring to the maximal order minors, we can introduce the following right-primeness notions and the analogous definitions of left-primeness are obvious.

An L-polynomial matrix $G \in F[z, z^{-1}]^{p \times m}$ is:

- right minor prime (rMP) if $p \geq m$ and all the L-polynomials in the ideal τ_G, generated by its maximal order minors, are devoid of (nontrivial) common factors.

- right variety prime (rVP) if $p \geq m$ and the ideal τ_G includes (nonzero) L-polynomials in $F[z_i, z_i^{-1}]$, for every $i = 1, 2, \cdots, n$.
- right zero prime (rZP) if $p \geq m$ and the ideal τ_G is the ring $F[z, z^{-1}]$ itself. In particular, when $p = m$, a right (and hence left) zero prime matrix is called unimodular.

From the above definition, It is obvious that in the 1-D case all properties of the above definition are equivalent to right factor primeness. For 2-D L-polynomial matrices, only factor, minor and variety primeness coincide. As a simple example in the 2-D case, consider the polynomials:

$$a(z_1, z_2) = z_1 - 1, b(z_1, z_2) = z_2 - 1$$

These two polynomials are not factor coprime but have common zeros.

In 2-D systems, the concepts of common divisor, greatest common divisor, common multiple, least common multiple can be defined as those in 1-D case with left and right being the two scenarios which depend on the location of the divisor (multiple).

For 2-D polynomial matrices, the three coprimeness are defined by:

- Polynomial matrices $A \in \mathbb{F}^{p\times m}[z_1, z_2], B \in \mathbb{F}^{q\times m}[z_1, z_2]$ are called Minor Right Coprime(MRC) if all the $m \times m$ minors of the matrix $\begin{bmatrix} A \\ B \end{bmatrix}$ are coprime(relatively prime), i.e. their greatest common polynomial divisor is a non-zero constant.
- Polynomial matrices $A \in \mathbb{F}^{p\times m}[z_1, z_2], B \in \mathbb{F}^{q\times m}[z_1, z_2]$ are called Factor Right(Left) Coprime (FRC/FLC) if great right (left) common divisor (GRCD/GLCD) is a unimodular matrix $U = U(z_1, z_2)(det(U) \in \mathbb{R})$.
- Polynomial matrices $A \in \mathbb{F}^{p\times m}[z_1, z_2], B \in \mathbb{F}^{q\times m}[z_1, z_2](p + q \geq m \geq 1)$ are called Zero Right Coprime(ZRC) if there exists no pair (z_1, z_2) which is a zero of all the $m\times m$ minors of the matrix $\begin{bmatrix} A \\ B \end{bmatrix}$. Polynomial matrices $A \in \mathbb{F}^{p\times m}[z_1, z_2], B \in \mathbb{F}^{q\times m}[z_1, z_2]$ are called Zero Left Coprime(ZLC) if the transposed matrices A^T, B^T are ZRC.

Consider two polynomial matrices $A \in \mathbb{F}^{p\times m}[z_1, z_2], B \in \mathbb{F}^{q\times m}[z_1, z_2](p + q \geq m \geq 1)$. The matrices is Zero Left Coprime(ZLC) if there exists no pair (z_1, z_2) which is a zero of all the $m \times m$ minors of the matrix $\left[\begin{array}{c|c} A & B \end{array}\right]$. A and B are ZLC if and if only there exists two polynomial $X = X(z_1, z_2) \in F^{p\times m}[z_1, z_2], Y = Y(z_1, z_2) \in F^{q\times m}[z_1, z_2]$ such that:

$$AX + BY = I_m$$

The primeness of polynomial matrices play a important role in the polynomial algebra and ZRC and ZLC often serve as the conditions of certain algorithms. In this thesis, the concept of ZRC is applied to determine the existence of dead-beat observer and residual generator.

2.2.2 Factorization

In 2-D systems, due to the two indeterminates z_1 and z_2, it is difficult to factorize directly for the general case. Let $\mathbb{F}[z_1][z_2]$ indicate the polynomial matrices with z_2 as terms and z_1 as the coefficients, vise versa. The primitive matrix $A(z_1, z_2)$ is a polynomial matrix which is full rank for all fixed z_1^0 in $\mathbb{F}[z_1][z_2]$. Similarly, $A(z_1, z_2)$ also holds for $\mathbb{F}[z_2][z_1]$. Let $A(z_1, z_2) \in \mathbb{F}^{m\times n}[z_1, z_2](m \leq n)$ be a full rank matrix, then there exists a unique $A^*(z_1, z_2)$ (modulo a left unimodular matrix) and a unique $\bar{A}(z_1, z_2)$ (modulo a right unimodular matrix) such that

$$A(z_1, z_2) = \bar{A}(z_1, z_2)A^*(z_1, z_2)$$

where $det(\bar{A})(z_1, z_2) = \bar{a}(z_2) \in \mathbb{F}[z_2]$ and $A^*(z_1, z_2)$ is primitive in $F[z_2][z_1]$. The primitive factorization of $A^*(z_1, z_2)$ is therefore obtained. The primitive factorization only extract one polynomial which does not contain one of the indeterminates, and another polynomial is still a primitive polynomial.

If $det(A(z_1, z_2)) = \prod_{i=1}^{k} a_i(z_1, z_2)$, a more general factorization result can be obtained and the polynomial matrix $A(z_1, z_2)$ can be factored such that:

$$A(z_1, z_2) = \prod_{i=1}^{k} A_i(z_1, z_2) \tag{2.1}$$

where $det(A_i(z_1, z_2)) = a_i(z_1, z_2)(i = 1, 2, \cdots, k)$ and $a_i(z_1, z_2)$ are arbitrary polynomials. Clearly, the factorization in general is not unique since the possibilities of many factored polynomial $a_i(z_1, z_2)$.

2.2.3 Matrix fraction description

In 2-D polynomial matrices, it is essential to obtain a factored matrix pair and then perform the related calculations. Consider a 2-D rational matrix(which means the matrix could be reducible) $G(z_1, z_2) \in \mathbb{F}^{m\times n}[z_1, z_2]$, it can always be written as:

$$G(z_1, z_2) = BA^{-1} \tag{2.2}$$

or

$$G(z_1, z_2) = A_1^{-1} B_1 \tag{2.3}$$

where $A = A(z_1, z_2) \in \mathbb{F}^{n\times n}[z_1, z_2]$,$B = B(z_1, z_2) \in \mathbb{F}^{m\times n}[z_1, z_2]$, and $A_1 = A_1(z_1, z_2) \in \mathbb{F}^{m\times m}[z_1, z_2]$,$B_1 = B_1(z_1, z_2) \in \mathbb{F}^{m\times n}[z_1, z_2]$.

The equations. (2.2) and (2.3) are called matrix fraction description(MFD). If B and A are right coprime, that is A, B have no nontrivial right common factor, then (2.2) is a irreducible representation. MFD requires to find the greatest common right divisor(GCRD) of polynomial matrices. Let

$$U \begin{bmatrix} A \\ \hline B \end{bmatrix} = \begin{bmatrix} R \\ \hline 0 \end{bmatrix}$$

where U, A, B, R are 2-D polynomials, and $det(U) \in \mathbb{F}[z_2]$. If R has a primitive left factorization $R = \bar{R}R^*$ in $\mathbb{F}[z_2][z_1]$, then R^* is a right common factor of A and B.

Let $G(z_1, z_2) \in \mathbb{F}^{l \times m}(z_1, z_2)$ be a 2-D system described as transfer function:

$$G(z_1, z_2) = \begin{bmatrix} G_{11}(z_1, z_2) & \cdots & G_{1m}(z_1, z_2) \\ \vdots & \vdots & \vdots \\ G_{l1}(z_1, z_2) & \cdots & G_{lm}(z_1, z_2) \end{bmatrix} \tag{2.4}$$

Using the algorithm (Morf et al, 1977) to find the least common denominator of G_{1k}, G_{2k} , $\cdots$, G_{mk} ($k = 1, 2, \cdots, m$). The above equation can be rewritten as:

$$G(z_1, z_2) = \begin{bmatrix} \frac{N_{11}(z_1, z_2)}{D_1(z_1, z_2)} & \cdots & \frac{N_{1m}(z_1, z_2)}{D_m(z_1, z_2)} \\ \vdots & \vdots & \vdots \\ \frac{N_{l1}(z_1, z_2)}{D_1(z_1, z_2)} & \cdots & \frac{N_{lm}(z_1, z_2)}{D_m(z_1, z_2)} \end{bmatrix} = N_R D_R^{-1} \tag{2.5}$$

where

$$N_R = \begin{bmatrix} N_{11}(z_1, z_2) & \cdots & N_{1m}(z_1, z_2) \\ \vdots & \vdots & \vdots \\ N_{l1}(z_1, z_2) & \cdots & N_{lm}(z_1, z_2) \end{bmatrix}$$

and

$$D_R = \begin{bmatrix} D_1(z_1, z_2) & 0 & \cdots & 0 \\ 0 & D_2(z_1, z_2) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & D_m(z_1, z_2) \end{bmatrix}$$

The MFD is thus achieved by the above equations.

2.2.4 Bézout identity

Let $N_R(z_1, z_2),\ D_R(z_1, z_2)$ (or $N_L(z_1, z_2), D_L(z_1, z_2)$) be two polynomial matrices being right (left) coprime, then there exists a polynomial matrix in z_2, say $E_R(z_2)$ (or $E_L(z_2)$) and two polynomial matrices $X_R,\ Y_R\ (X_L,\ Y_L)$ such that:

$$\begin{aligned} X_R(z_1,z_2)N_R(z_1,z_2)+Y_R(z_1,z_2)D_R(z_1,z_2)&=E_R(z_2)\\ N_L(z_1,z_2)Y_L(z_1,z_2)-D_L(z_1,z_2)X_L(z_1,z_2)&=E_L(z_2) \end{aligned} \tag{2.6}$$

Let $G(z_1,z_2)=N_R D_R^{-1}=D_L^{-1}N_L$ with $N_R=N_R(z_1,z_2), D_R=D_R(z_1,z_2)$ right coprime and $N_L=N_L(z_1,z_2), D_L=D_L(z_1,z_2)$ left coprime, then $\det(D_L)=\det(D_R)$. The equation (2.6) is usually called Bézout identity in 2-D systems. The problem of solving the Bézout identity equation needs to extend the solution to a general case:

$$AX+BY=C$$

or

$$XA+YB=C$$

where $A\in\mathbb{R}^{p\times l}[z_1,z_2], B\in\mathbb{R}^{q\times l}[z_1,z_2], C\in\mathbb{R}^{m\times l}[z_1,z_2], X\in\mathbb{R}^{m\times p}[z_1,z_2], Y\in\mathbb{R}^{m\times q}[z_1,z_2]$.

The necessary and sufficient condition for the existence of solution of general polynomial matrices equations is expressed as:

$$\bar{A}\bar{X}+\bar{B}\bar{Y}=\bar{C} \tag{2.7}$$

where $\bar{A}=\begin{bmatrix}A&0\\0&0\end{bmatrix}\in\mathbb{R}^{r\times r}[z_1,z_2]$, $\bar{B}=\begin{bmatrix}B\\0\end{bmatrix}\in\mathbb{R}^{r\times q}[z_1,z_2]$, $\bar{C}=\begin{bmatrix}C\\0\end{bmatrix}\in\mathbb{R}^{r\times m}[z_1,z_2]$, $\bar{X}=\begin{bmatrix}X\\X'\end{bmatrix}\in\mathbb{R}^{r\times m}[z_1,z_2]$, $\bar{Y}=Y\in\mathbb{R}^{q\times m}[z_1,z_2]$ and $r\in\max(l,p+q)$,$X'\in\mathbb{R}^{(r-p)\times m}[z_1,z_2]$ is arbitrary.

If equation (2.7) has a solution, then the general case also has a solution. There always exists a unimodular matrix $U\in\mathbb{R}^{(r+q)\times(r+q)}[z_1,z_2]$ such that:

$$\left[\begin{array}{c|c}\bar{A}&\bar{B}\end{array}\right]U=\left[\begin{array}{c|c}G&0\end{array}\right] \tag{2.8}$$

where $G \in \mathbb{R}^{r \times r}[z_1, z_2]$ and G is the greatest common left divisor (GCLD) of $\bar{A}$ an $\bar{B}$.

The algorithm of finding the solution of $AX + BY = C$ could be achieved by applying the elementary column operation to carry out the reduction:

$$\begin{bmatrix} A & B & C \\ I_l & 0 & 0 \\ 0 & I_q & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} A & B & 0 \\ I_l & 0 & U_1 \\ 0 & I_q & U_2 \end{bmatrix}$$

So the solution would be $X = -U_1$ and $Y = -U_2$. Also, the solution of $XA + YB = C$ could be found by utilizing the transpose of $AX + BY = C$.

Another algorithm is given by the solution of $\bar{A}\bar{X} + \bar{B}\bar{Y} = \bar{C}$. First, by defining the equation (2.8)'s $\bar{A}$, $\bar{B}$, $\bar{C}$ and $r = \max(l, p + q)$. Then, using the elementary column operation (the elementary column operation could be found in literature (Galkowski, 2001)) to carry out the following reduction, it is easy to obtain:

$$\begin{bmatrix} \bar{A} & \bar{B} \\ I_l & 0 \\ 0 & I_q \end{bmatrix} \longrightarrow \begin{bmatrix} G & 0 \\ U_1 & U_2 \\ U_3 & U_4 \end{bmatrix}$$

then the solution $\bar{X}$ and $\bar{Y}$ can be obtained by equation:

$$\begin{bmatrix} \bar{X} \\ \bar{Y} \end{bmatrix} = \begin{bmatrix} U_1 & U_2 \\ U_3 & U_4 \end{bmatrix} \begin{bmatrix} \bar{C}_o \\ T \end{bmatrix}$$

with $\bar{C} = G\bar{C}_0$ and T is arbitrary. Finally, the solution of X and Y could be obtained by the definitions of $\bar{X}$ and $\bar{Y}$.

2.2.5 2-D transfer function and realization

Similar to 1-D systems described in section 1, a 2-D transfer function can be expressed by a combination of system matrices and the conversion process is called realization. For Roesser model (1.1), its transfer function can be written as:

$$G(z_1, z_2) = C(Z - A)^{-1}B + D$$

where $Z = \begin{bmatrix} I_1 z_1 & 0 \\ 0 & I_2 z_2 \end{bmatrix}$. The transfer function of F-M model (1.2) is

$$G(z_1, z_2) = C(I_n z_1 z_2 - A_1 z_1 - A_2 z_2)^{-1}(B_1 z_1 + B_2 z_2) + D$$

The transfer function of Roesser model can also be written in the form

$$G(z_1, z_2) = C(z_2)[I_{n_1} z_1 - A(z_2)]^{-1}B(z_2) + D(z_2) \tag{2.9}$$

where

$$A(z_2) = A_{11} + A_{12}[I_{n_2} z_2 - A_{22}]^{-1}A_{21}$$

$$B(z_2) = B_1 + A_{12}[I_{n_2} z_2 - A_{22}]^{-1}B_2$$

$$C(z_2) = C_1 + C_2[I_{n_2} z_2 - A_{22}]^{-1}A_{21}$$

$$D(z_2) = C_2[I_{n_2} z_2 - A_{22}]^{-1}B_2 + D$$

The equation (2.9) is called first level realization. The realization algorithm for F-M model is present by Fornasini and Maichesini (1976).

In some recent research, the transfer functions of F-M model and Roesser model have been written differently due to different definition of the Z-transform. By defining (Bisiacco et al., 1989)

$$\begin{aligned} X(z_1, z_2) &= \sum_{i+j>0} x(i,j) z_1^i z_2^j \\ U(z_1, z_2) &= \sum_{i+j>0} u(i,j) z_1^i z_2^j \\ Y(z_1, z_2) &= \sum_{i+j>0} y(i,j) z_1^i z_2^j \end{aligned} \tag{2.10}$$

the transfer function of F-M model (1.2) can be obtained:

$$G(z_1, z_2) = C(I_n - A_1 z_1 - A_2 z_2)^{-1}(B_1 z_1 + B_2 z_2) + D \tag{2.11}$$

The transfer function of Roesser model becomes:

$$G(z_1, z_2) = CZ(I - AZ)^{-1}B + D \tag{2.12}$$

In the observer based fault detection method described in the next Chapter, the transfer function of Roesser model in 2.12 is used.

2.3 Polynomial matrices applications in fault detection of 2-D systems

Polynomial matrices are useful in representing the system dynamics, such as the transfer functions, the observability of 2-D systems. In this thesis, most applications of polynomial matrices are for developing the observer based fault detection method. In the observer based fault detection, in order for an observer and residual generator to exist, zero primeness of the corresponding polynomial matrices has to be satisfied, meaning that the solution of Equation 2.1 is a pair of non-zero polynomial matrices. The Bézout identity presented in 2.2.4 is necessary in the observer based fault detection method as it is required in deriving the transfer functions of the observers and residual generator.

The concept of minimal left annihilator (MLA) (Rocha and Willems, 1991) is the 2-D expansion of zero space involved with solution of polynomial equations. For a given polynomial $R(z_1, z_2)$, if R is a $p \times q$ polynomial matrix of rank r, a polynomial matrix L is a left annihilator of R if $LR = 0$. A left annihilator L of R is an MLA if it is of full row rank and for any other left annihilator M of R, we have $M = PL$ for some polynomial matrix P. If $p = r$, the MLA of R is clearly an empty matrix, which means a zero-dimensional matrix. By this definition, zero space of 2-D system is achieved. MLA is always exists with the $(p - r) \times p$ left factor prime. Calculating MLA is also required in deriving the observer based fault detection.

Realization is necessary to derive the observers and residual generators in the observer based fault detection method. 2-D polynomial matrices are used for realization related calculations. With the obtained 2-D transfer functions, MFD strategies in section 2.2.3 can be used to decompose the transfer function to two polynomial matrices corresponding to its denominator and numerator. Polynomial matrix equations are also utilized in the procedure of forming the system matrices.

Chapter 3

Observer Based FDI for Roesser Model in Two Dimensional Systems

3.1 Introduction

Fault detection is an integral component of system engineering. There is an abundance of literature on process fault detection ranging from analytical methods to artificial intelligence and statistical approaches (Wu et al., 2009; Yao et al., 2011). A priori process knowledge plays a crucial role in diagnostic decision-making. A priori knowledge could be in the form of material and energy balances (Mehra and Peschon, 1971; Kramer, 1987), suitably transformed process models (Willsky and Jones, 1976; Gertler, 1991; Frank, 1990) or distribution information in the measurement space (Hoskins and Himmelblau, 1988). The models that have been heavily investigated in fault diagnosis are input-output or state space models. Faults in the state space framework are described in a series of papers by (Gertler, 1991, 1993) and in recent books on fault detection and diagnosis (Gertler, 1998; Chen and Patton, 1999; Russell et al., 2000). Development of fault detection and diagnosis techniques have been predominantly focused on 1-D system.

Fault detection for 2-D state-space models has been considered by only a few researchers, e.g. in (Fornasini and Maichesini, 1988). In these papers, the parity check relations were explored for 2-D state-space models by means of a polynomial approach. The dead-beat observer has been introduced for state estimate for decades (Zampieri, 1991; Kaczorek, 2001). In the recent years, the dead-beat observer was introduced into fault detection of processes with the F-M model (Bisiaco and Valcher, 2006; Bisiacco and Valcher, 2008). Very few research has been carried out on fault detection for systems by the Roesser model (Wu et al., 2011, 2012; Maleki et al., 2013).

In this chapter, an observer based fault detection method is investigated for systems described by the Roesser model (Wang and Shang, 2014). From the Roesser model, the PBH observability matrix is constructed and examined, and the transfer function of a dead beat observer is then developed by solving the Bézout equation of the PBH observability matrix. The observer is obtained from its transfer function using an efficient realization technique so that the resulting observer is of minimal order. Fault detection is performed through monitoring the residual, which is closely correlated with the fault signal. With the state estimate from the observer as one of its input, a residual generator is designed such that its output residual signal reflects occurrence of faults. Research is carried out on fault detection for systems without disturbances as well as those affected by unknown disturbances. Development of fault detection for systems affected by unknown disturbances requires stricter conditions for the existence of the observer and residual generator. Efficiency of the fault detection is ensured by using an enhanced realization technique and the strategy of minimizing delay in residual generator. The proposed fault detection is shown to be effective in efficiently detecting occurrence of faults, although unknown disturbances may sometimes jeopardize the performance of the strategy.

3.2 Realization of Roesser model

A Roesser model can be written as:

$$
\begin{aligned}
\begin{bmatrix} x^h(i+1,j) \\ x^v(i,j+1) \end{bmatrix} &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x^h(i,j) \\ x^v(i,j) \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u(i,j), \\
y(i,j) &= \begin{bmatrix} C_1 & C_2 \end{bmatrix} \begin{bmatrix} x^h(i,j) \\ x^v(i,j) \end{bmatrix} + Ku(i,j),
\end{aligned}
\tag{3.1}
$$

where i and j indicate the horizontal (or spacial) and vertical (or time) indices, x^h is the state variable evolving horizontally (or spatially), x^v is the state variable evolving vertically (or in time), $u(i,j)$ and $y(i,j)$ indicate the process input and output, respectively, A's, B's, C's and K are the real matrices of suitable dimensions. The system can be denoted in abbreviation by

$$\{A_{11}, A_{12}, A_{21}, A_{22}, B_1, B_2, C_1, C_2, K\}.$$

For simplicity of description, the following matrices are defined:

$$
\begin{aligned}
A &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad Z = \begin{bmatrix} z_1 I_1 & \mathbf{0} \\ \mathbf{0} & z_2 I_2 \end{bmatrix}, \quad x = \begin{bmatrix} x^h \\ x^v \end{bmatrix}, \\
B &= \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad C = \begin{bmatrix} C_1 & C_2 \end{bmatrix},
\end{aligned}
\tag{3.2}
$$

where I_1 and I_2 are two unit matrices with the dimension of A_{11} and A_{22}, respectively. The system in (3.1) is therefore denoted by $\{A, B, C, K\}$. In the following description, where necessary, the subscript is used to distinguish the unit matrices with different dimensions, e.g., I_x indicates the unit matrix with the dimension of $x(i,j)$, I_{x+y} has the dimension of $x(i,j)$ plus that of $y(i,j)$, etc.

The transfer functions from input to output can be obtained as:

$$W(z_1, z_2) = CZ(I - AZ)^{-1}B + K. \tag{3.3}$$

In the proposed fault detection method, it is necessary to construct the state space Roesser models from a 2-D transfer function. An efficient realization technique is essential in order for the fault detection method to be implementable. Realization of a 2-D transfer function polynomial is more complicated due to the two dimensional structure. The state space realizations may have many variations such as normal form, Jordan form or canonical form. In the proposed observer based fault detection, realization of state observers and residual generators are required. It is desirable to generate system matrices of minimal order with as many sparse matrices in it as possible such that the state space models have simple structure requiring minimal computations. In the 2-D Roesser model in (1.1), there exist horizontal and vertical state variables and the system matrices have to be produced and decomposed to fit the structure of $\{A_{11}, A_{12}, A_{21}, A_{22}, B_1, B_2, C_1, C_2, K\}$.

In this chapter, a realization method with minimal order is used to construct the Roesser models (Xu et al., 2008). This procedure can be carried out with the following steps:

Step 1. From (3.1) and (3.3), the matrix K can be easily obtained as $W(0,0) = K$.

Step 2. Denote the remaining part of the polynomial $W(z_1, z_2)$ as $W_1(z_1, z_2)$, i.e., $W_1(z_1, z_2) = W(z_1, z_2) - W(0,0)$. Decompose the polynomial $W_1(z_1, z_2)$ into $W_1(z_1, z_2) = N(z_1, z_2)D(z_1, z_2)^{-1}$ using the matrix fraction description method. Without loss of generality, it is assumed that $D(0,0) = I$.

Step 3. Let $D_T(z_1, z_2) = I - D(z_1, z_2)$ and form the following polynomial:

$$F(z_1, z_2) = \begin{bmatrix} N(z_1, z_2) \\ D_T(z_1, z_2) \end{bmatrix}. \tag{3.4}$$

Step 4. For each column of $F(z_1, z_2)$, construct column vectors $\tilde{\phi}_{1i}$ consisting of the monomial entries from the k-th column of $F(z_1, z_2)$ in the descending order as $\tilde{\phi}_{1k} = \begin{bmatrix} z_1^{r_{1k}} & z_1^{r_{1k}-1} & \cdots & z_1 \end{bmatrix}^T$ and $\tilde{\phi}_{2k}$ as $\tilde{\phi}_{2k} = \begin{bmatrix} z_2^{r_{2k}} & z_2^{r_{2k}-1} & \cdots & z_2 \end{bmatrix}^T$, where $k = 1, \cdots, l$, l indicates the number of columns in $F(z_1, z_2)$, r_{1k} is the highest order of the z_1 terms in the k-th column in $F(z_1, z_2)$, and r_{2k} is the highest order of the z_2 terms in the k-th column in $F(z_1, z_2)$. Insert the power product terms in the k-th column of $F(z_1, z_2)$ into either $\tilde{\phi}_{1k}$ or $\tilde{\phi}_{2k}$. Additional terms can then be added if necessary such that for each term other than z_1 in $\tilde{\phi}_{1k}$ there is another term satisfying $\tilde{\phi}_{1k}(k_1) = z_1\tilde{\phi}_{1k}(j_1)$, and for each term other than z_2 in $\tilde{\phi}_{2k}$ there exists another term satisfying $\tilde{\phi}_{2k}(k_2) = z_2\tilde{\phi}_{2k}(j_2)$. Let $\phi_{1k} = z_1^{-1}\tilde{\phi}_{1k}$ if $\tilde{\phi}_{1k}$ is not empty, and $\phi_{2k} = z_2^{-1}\tilde{\phi}_{2k}$ if $\tilde{\phi}_{2k}$ is not empty. The matrix Φ can then be formed:

$$\Phi = \begin{bmatrix} \phi_{11} & & \\ & \ddots & \\ & & \phi_{1l} \\ \phi_{21} & & \\ & \ddots & \\ & & \phi_{2l} \end{bmatrix}. \tag{3.5}$$

Step 5. Find D_H and N_H such that $D_T = D_H Z\Phi$ and $N = N_H Z\Phi$.

Step 6. Construct A_0 and B such that

$$\Phi = (I - A_0)\Phi + B, \tag{3.6}$$

for some matrices A_0 and B.

Step 7. With matrices A_0 and B satisfying (3.6), the matrices A and C in the Roesser model are obtained:

$$\begin{aligned} A &= A_0 + BD_{HT}, \\ C &= N_{HT}. \end{aligned} \tag{3.7}$$

With the $\{A, B, C, K\}$, a Roesser model is then obtained.

In the next section, observer based fault detection for Roesser models is developed and the realization technique is used to derive 2-D state space models for fault detection. It is necessary to form the state space models of minimal order by ensuring that the matrix Φ is of minimal dimension. In the above algorithm, ϕ_i is constructed such that it contains no redundant terms, and thus can be used to derive the Roesser model for efficient fault detection.

3.3 Observer-based fault detection for the Roesser model

Consider a Roesser model with fault signals:

$$\begin{aligned}
\begin{bmatrix} x^h(i+1,j) \\ x^v(i,j+1) \end{bmatrix} &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x^h(i,j) \\ x^v(i,j) \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u(i,j) + \begin{bmatrix} F_1 \\ F_2 \end{bmatrix} f(i,j), \\
y(i,j) &= \left[\begin{array}{cc} C_1 & C_2 \end{array} \right] \begin{bmatrix} x^h(i,j) \\ x^v(i,j) \end{bmatrix} + K_u u(i,j) + K_f f(i,j),
\end{aligned} \tag{3.8}$$

where $f(i,j)$ indicates the fault signal. Let

$$F = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix}. \tag{3.9}$$

The system in (3.8) is denoted in abbreviation by $\{A, B, F, C, K_u, K_f\}$. The transfer functions for input to output and fault to output can be expressed as:

$$\begin{aligned}
W_u(z_1, z_2) &= CZ(I - AZ)^{-1}B + K_u, \\
W_f(z_1, z_2) &= CZ(I - AZ)^{-1}F + K_f.
\end{aligned} \tag{3.10}$$

In developing a fault detection strategy, a dead-beat observer is first constructed for a Roesser model. The state estimate obtained from the observer is then input to a fault detection residual generator. The structure of the observer based fault detection is as shown in Fig. 3.1. Note that development of fault detection requires a state observer and a residual generator.

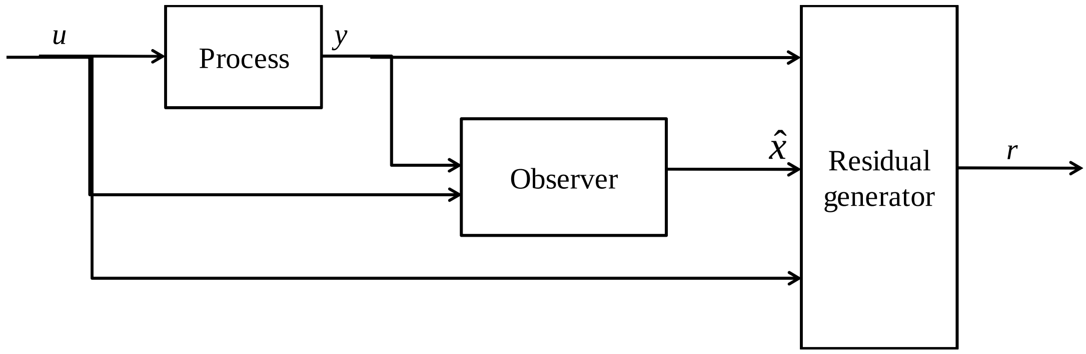


Figure 3.1: Structure of the observer based fault detection.

3.3.1 Observer development

In constructing a state space observer for the system in (3.8), the observability PBH matrix $\sigma(z_1, z_2)$ can be constructed:

$$\sigma(z_1, z_2) = \begin{bmatrix} I - ZA \\ C \end{bmatrix}. \tag{3.11}$$

Note that the PBH matrix is a polynomial matrix. For constructing an observer, it is necessary that the polynomial matrix is zero right prime (ZRP). An observer transfer function exists if and only if the polynomial matrix is ZRP. ZRP of the PBH matrix can be determined by examining the Bezout equation of $\sigma(z_1, z_2)$:

$$\begin{bmatrix} Q(z_1, z_2) & P(z_1, z_2) \end{bmatrix} \sigma(z_1, z_2) = \begin{bmatrix} Q(z_1, z_2) & P(z_1, z_2) \end{bmatrix} \begin{bmatrix} I_x - ZA \\ C \end{bmatrix} = I_x. \tag{3.12}$$

If the Bézout equation is solvable, $\sigma(z_1, z_2)$ is ZRP and there exists a dead beat observer. The observer has process input $u(i, j)$ and output $y(i, j)$ as its input and generates a state estimate $\hat{x}(i, j)$ as its output.

Theorem 1. Given the polynomial pair Q and P satisfying (3.12), an observer of the system in

(3.8) can be designed with its transfer function being:

$$\begin{aligned} \hat{W}_o(z_1, z_2) &= \left[\begin{array}{cc} \hat{W}_u(z_1, z_2) & \hat{W}_y(z_1, z_2) \end{array} \right] \\ &= \left[\begin{array}{cc} Q(z_1, z_2)ZB - P(z_1, z_2)K_u & P(z_1, z_2) \end{array} \right]. \end{aligned} \tag{3.13}$$

Proof: When there is no fault signal, i.e., $f(i, j) = 0$, the transfer function of the Roesser model in (3.8) can be written as:

$$W_u(z_1, z_2) = C(I - ZA)^{-1}ZB + K_u, \tag{3.14}$$

which is equivalent to (3.10). For convenience of description, $u(z_1, z_2)$, $y(z_1, z_2)$, $x(z_1, z_2)$ and $\hat{x}(z_1, z_2)$ are used to indicate the z-transform of the input $u(i, j)$, output $y(i, j)$, state $x(i, j)$ and state estimate $\hat{x}(i, j)$, respectively. The output $y(z_1, z_2)$ is then expressed as:

$$y(z_1, z_2) = C(I - ZA)^{-1}x_0 + (C(I - ZA)^{-1}ZB + K_u)u(z_1, z_2), \tag{3.15}$$

where x_0 indicates the initial value of the state $x(i, j)$. From (3.13), the transfer function of the observer is comprised of $\hat{W}_u$ and $\hat{W}_y$:

$$\begin{aligned} &\hat{W}_u(z_1, z_2) = Q(z_1, z_2)ZB - P(z_1, z_2)K_u, \\ &\hat{W}_y(z_1, z_2) = P(z_1, z_2). \end{aligned} \tag{3.16}$$

Denote the state space form of the observer as $\Sigma = \left\{ \hat{A},\ \hat{B},\ \hat{C},\ \hat{K} \right\}$. The state x and state estimate $\hat{x}$ can be described as:

$$\begin{aligned} &x(z_1, z_2) = (I - ZA)^{-1}x_0 + (I - ZA)^{-1}ZBu(z_1, z_2), \\ &\hat{x}(z_1, z_2) = \hat{C}(I - Z\hat{A})^{-1}\hat{x}_0 + \hat{W}_u(z_1, z_2)u(z_1, z_2) + \hat{W}_y(z_1, z_2)y(z_1, z_2), \end{aligned} \tag{3.17}$$

where $\hat{x}_0$ indicate the initial value of the state estimate $\hat{x}(i,j)$. By substituting (3.15) into (3.17), the difference between x and $\hat{x}$ is obtained:

$$\begin{aligned} E(z_1,z_2) &= x(z_1,z_2) - \hat{x}(z_1,z_2) \\ &= [(I-ZA)^{-1} - \hat{W}_y C(I-ZA)^{-1}]x_0 - \hat{C}(I-Z\hat{A})^{-1}\hat{x}_0 \\ &+ [(I-ZA)^{-1}ZB - \hat{W}_u(z_1,z_2) - \hat{W}_y(z_1,z_2)C(I-ZA)^{-1}ZB - \hat{W}_y(z_1,z_2)K_u] \\ &u(z_1,z_2). \end{aligned} \tag{3.18}$$

From (3.12), the polynomial matrix $Q(z_1,z_2)$ is:

$$Q(z_1,z_2) = (I-ZA)^{-1} - P(z_1,z_2)C(I-ZA)^{-1}. \tag{3.19}$$

Substituting (3.16) and (3.19) into (3.18) yields:

$$E(z_1,z_2) = Q(z_1,z_2)x_0 - \hat{C}(I-Z\hat{A})^{-1}\hat{x}_0. \tag{3.20}$$

It is noted that when an observer is designed with the transfer function as in (3.13), $E(z_1,z_2)$ is only affected by the initial conditions and the polynomial matrices. This prove the theorem. ■

When there exists a fault signal, i.e., $f(i,j) \neq 0$, it can be obtained:

$$\begin{aligned} E(z_1,z_2) &= Q(z_1,z_2)x_0 - \hat{C}(I-Z\hat{A})^{-1}\hat{x}_0 \\ &\quad + [Q(z_1,z_2)ZF - P(z_1,z_2)K_f]\, f(z_1,z_2). \end{aligned} \tag{3.21}$$

Equations (3.20) and (3.21) indicate that the initial values of the state x and state estimate $\hat{x}$ affect the performance of an observer. For simplicity of calculation, initial conditions of zeros are used in this chapter.

The designed observer requires to calculate the polynomial matrices $P(z_1,z_2)$ and $Q(z_1,z_2)$. From the Bézout equation of PBH matrix, $P(z_1,z_2)$ and $Q(z_1,z_2)$ can be calculated as a left inverse of the PBH matrix. The Polynomial Toolbox Software was used to calculate the inverse

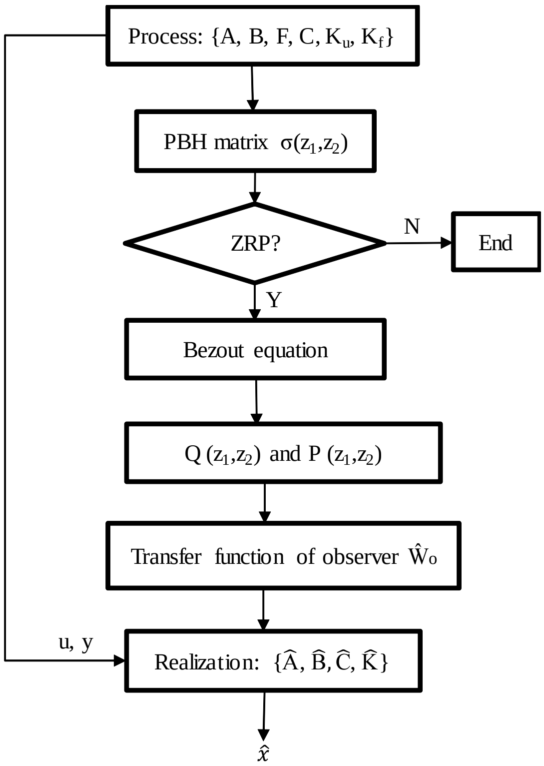


Figure 3.2: Construction of a state observer for a 2-D Roesser system.

of polynomial matrices in this work. With the resulting $P(z_1, z_2)$ and $Q(z_1, z_2)$, the transfer function of the observer is then obtained using Equation (3.13). In calculating the observer, it is essential to ensure that the Bézout equation is solvable, i.e., the PBH is zero right prime.

From the obtained transfer function (3.13), the realization technique described in the last section can be used to obtain a state space model of the observer, from which the evolvement of state estimate $\hat{x}$ is determined. The state space model of the observer takes the following form:

$$\begin{gathered}\begin{bmatrix} w^h(i+1,j) \\ w^v(i,j+1) \end{bmatrix} = \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22} \end{bmatrix} \begin{bmatrix} w^h(i,j) \\ w^v(i,j) \end{bmatrix} + \begin{bmatrix} \hat{B}_1 \\ \hat{B}_2 \end{bmatrix} \begin{bmatrix} u(i,j) \\ y(i,j) \end{bmatrix}, \\ \hat{x}(i,j) = \begin{bmatrix} \hat{C}_1 & \hat{C}_2 \end{bmatrix} \begin{bmatrix} w^h(i,j) \\ w^v(i,j) \end{bmatrix} + \hat{K} \begin{bmatrix} u(i,j) \\ y(i,j) \end{bmatrix}. \end{gathered} \tag{3.22}$$

From the state space observer, evolvement of the state can be estimated and the state estimate $\hat{x}$ is obtained. The schematic of observer construction is shown in Fig. 3.2.

3.3.2 Residual generator for fault detection

In the proposed method, faults are detected using a residual generator with its output being a residual signal. An effective residual generator is such that the output residual signal is zero if no fault occurs and is nonzero when a fault affects the system. The designed residual generator uses three types of input signals: process input $u(i,j)$, output $y(i,j)$, and the state estimate $\hat{x}(i,j)$ from the observer. Corresponding to these signals, the transfer function of the residual generator is comprised of three components:

$$\bar{W}_r(z_1,z_2) = \begin{bmatrix} \bar{W}_{ru}(z_1,z_2) & \bar{W}_{ry}(z_1,z_2) & \bar{W}_{r\hat{x}}(z_1,z_2) \end{bmatrix}. \tag{3.23}$$

Denoting the minimum left annihilator (MLA) of the PBH matrix σ as the polynomial matrices pair $-N(z_1,z_2)$ and $M(z_1,z_2)$, i.e,

$$\begin{bmatrix} -N(z_1,z_2) & M(z_1,z_2) \end{bmatrix} \begin{bmatrix} I_x - ZA \\ C \end{bmatrix} = \mathbf{0}_x. \tag{3.24}$$

A fault-related polynomial matrix can be constructed as:

$$\Gamma_f(z_1,z_2) = \begin{bmatrix} N(z_1,z_2) & M(z_1,z_2) \end{bmatrix} \begin{bmatrix} ZF \\ K_f \end{bmatrix}. \tag{3.25}$$

There exists a residual generator relating fault signals with residuals if $\Gamma_f(z_1,z_2)$ is zero right prime. Based on the obtained matrices $Q(z_1,z_2)$, $P(z_1,z_2)$, $N(z_1,z_2)$ and $M(z_1,z_2)$, the following matrix equation is established:

$$\begin{bmatrix} Q(z_1,z_2) & P(z_1,z_2) \\ -N(z_1,z_2) & M(z_1,z_2) \end{bmatrix} \begin{bmatrix} I_x - ZA & -L(z_1,z_2) \\ C & S(z_1,z_2) \end{bmatrix} = I_{x+y}, \tag{3.26}$$

and $L(z_1,z_2)$ and $S(z_1,z_2)$ satisfying (3.26) can then be solved.

Let $\Lambda(z_1,z_2)$ denote an L-polynomial left inverse of $\Gamma_f(z_1,z_2)$, i.e., there exist $n_1 > 0$ and

$n_2 > 0$ such that

$$\Lambda(z_1, z_2)\Gamma_f(z_1, z_2) = z_1^{n_1} z_2^{n_2} I_f. \tag{3.27}$$

Theorem 2. Given the polynomial matrices $Q(z_1, z_2)$, $P(z_1, z_2)$, $N(z_1, z_2)$, $M(z_1, z_2)$, $L(z_1, z_2)$ and $S(z_1, z_2)$ satisfying (3.26), a residual generator can be designed with its transfer functions $\begin{bmatrix} \bar{W}_{ru}(z_1, z_2) & \bar{W}_{ry}(z_1, z_2) & \bar{W}_{r\hat{x}}(z_1, z_2) \end{bmatrix}$ being:

$$\begin{bmatrix} V(z_1, z_2) & \bar{W}_{ry}(z_1, z_2) \end{bmatrix} \begin{bmatrix} L(z_1, z_2) \\ S(z_1, z_2) \end{bmatrix} = \Lambda(z_1, z_2), \tag{3.28}$$

and

$$\begin{bmatrix} \bar{W}_{r\hat{x}}(z_1, z_2) & \bar{W}_{ru}(z_1, z_2) \end{bmatrix} = \begin{bmatrix} V(z_1, z_2) & -\bar{W}_{ry}(z_1, z_2) \end{bmatrix} \begin{bmatrix} I - ZA & -ZB \\ C & K_u \end{bmatrix}. \tag{3.29}$$

Proof: Input variables of a residual generator includes the process input $u(i, j)$, output $y(i, j)$ and state estimate $\hat{x}(i, j)$. The output $y(i, j)$ and state estimate $\hat{x}(i, j)$ are expressed as

$$\begin{aligned} y(z_1, z_2) &= C(I - ZA)^{-1}x_0 + (C(I - ZA)^{-1}ZB + K_u)u(z_1, z_2) \\ &\quad +(C(I - ZA)^{-1}ZF + K_f)f(z_1, z_2), \\ \hat{x}(z_1, z_2) &= \hat{C}(I - Z\hat{A})^{-1}\hat{x}_0 + \hat{W}_u(z_1, z_2)u(z_1, z_2) + \hat{W}_y(z_1, z_2)y(z_1, z_2) \\ &= \hat{C}(I - Z\hat{A})^{-1}\hat{x}_0 + \left[Q(z_1, z_2)ZB - P(z_1, z_2)K_u\right] u(z_1, z_2) \\ &\quad +P(z_1, z_2)y(z_1, z_2), \end{aligned} \tag{3.30}$$

where $f(z_1, z_2)$ indicates the z-transform of the fault signal $f(i, j)$. Denote the state space form of the residual generator as $\Sigma = \left\{\bar{A},\ \bar{B},\ \bar{C},\ \bar{K}\right\}$. The z-transform of the residual signal can then be written as:

$$
\begin{aligned}
r(z_1,z_2) =& \bar{W}_{ru}(z_1,z_2)u(z_1,z_2)+\bar{W}_{ry}(z_1,z_2)y(z_1,z_2) \\
&+\bar{W}_{r\hat{x}}(z_1,z_2)\hat{x}(z_1,z_2)+\bar{C}(I-Z\bar{A})^{-1}r_0,
\end{aligned} \tag{3.31}
$$

where r_0 is the initial value of the residual signal $r(i,j)$.

Substituting (3.30) to (3.31) and rearranging the equation yields

$$
\begin{aligned}
r(z_1,z_2) = & \left[(\bar{W}_{r\hat{x}}(z_1,z_2)+\bar{W}_{ry}(z_1,z_2)C)(I-ZA)^{-1}ZB+\bar{W}_{ru}(z_1,z_2)+\bar{W}_{ry}(z_1,z_2)K_u\right] \\
& u(z_1,z_2)+\left[\bar{W}_{r\hat{x}}(z_1,z_2)P(z_1,z_2)+\bar{W}_{ry}(z_1,z_2)\right] \\
& (C(I-ZA)^{-1}ZF+K_f)f(z_1,z_2)+\bar{C}(I-Z\bar{A})^{-1}r_0+ \\
& \bar{W}_{r\hat{x}}(z_1,z_2)\hat{C}(I-Z\hat{A})^{-1}\hat{x}_0+ \\
& \left[\bar{W}_{r\hat{x}}(z_1,z_2)P(z_1,z_2)+\bar{W}_{ry}(z_1,z_2)\right]C(I-ZA)^{-1}x_0.
\end{aligned} \tag{3.32}
$$

It is noted that the above expression is composed of the effects of input $u(z_1,z_2)$, fault $f(z_1,z_2)$ and initial conditions. Define

$$
\begin{aligned}
r_u(z_1,z_2) = & \left[\bar{W}_{r\hat{x}}(z_1,z_2)+\bar{W}_{ry}(z_1,z_2)C\right](I-ZA)^{-1}ZB \\
& +\bar{W}_{ru}(z_1,z_2)+\bar{W}_{ry}(z_1,z_2)K_u,
\end{aligned} \tag{3.33}
$$

$$
\begin{aligned}
r_f(z_1,z_2) = & \left[\bar{W}_{r\hat{x}}(z_1,z_2)P(z_1,z_2)+\bar{W}_{ry}(z_1,z_2)\right] \\
& (C(I-ZA)^{-1}ZF+K_f).
\end{aligned} \tag{3.34}
$$

It can be seen that $r_u(z_1,z_2)$ and $r_f(z_1,z_2)$ describes the effects of the input $u(z_1,z_2)$ and fault $f(z_1,z_2)$, respectively. From the transfer function of the designed residual generator in (3.29), $\bar{W}_{r\hat{x}}(z_1,z_2)$ and $\bar{W}_{ru}(z_1,z_2)$ can be described as:

$$
\begin{aligned}
\bar{W}_{r\hat{x}}(z_1,z_2) &= V(z_1,z_2)(I-ZA)-\bar{W}_{ry}(z_1,z_2)C, \\
\bar{W}_{ru}(z_1,z_2) &= -V(z_1,z_2)ZB-\bar{W}_{ry}(z_1,z_2)K_u.
\end{aligned} \tag{3.35}
$$

By substituting (3.35) to (3.33), it can be obtained

$$r_u(z_1, z_2) = 0. \tag{3.36}$$

The expression of $r_f(z_1, z_2)$ is then examined by substituting (3.35) to (3.34):

$$\begin{aligned} r_f(z_1, z_2) = &\big[V(z_1, z_2)(I - ZA)P(z_1, z_2) + \bar{W}_{ry}(I - CP(z_1, z_2))\big] \\ &(C(I - ZA)^{-1}ZF + K_f). \end{aligned} \tag{3.37}$$

Equation (3.26) is rewritten as

$$\begin{bmatrix} I_x - ZA & -L(z_1, z_2) \\ C & S(z_1, z_2) \end{bmatrix} \begin{bmatrix} Q(z_1, z_2) & P(z_1, z_2) \\ -N(z_1, z_2) & M(z_1, z_2) \end{bmatrix} = I_{x+y}, \tag{3.38}$$

which leads to the following equations:

$$\begin{aligned} &(I_x - ZA)P(z_1, z_2) = L(z_1, z_2)M(z_1, z_2), \\ &CP(z_1, z_2) = I_y - S(z_1, z_2)M(z_1, z_2). \end{aligned} \tag{3.39}$$

From (3.24), $C(I - ZA)^{-1} = M(z_1, z_2)^{-1}N(z_1, z_2)$. By substituting (3.39) to (3.37), Equation (3.37) becomes:

$$\begin{aligned} r_f(z_1, z_2) &= \big[V(z_1, z_2)L(z_1, z_2)M(z_1, z_2) + \bar{W}_{ry}S(z_1, z_2)M(z_1, z_2)\big] \\ &\quad M(z_1, z_2)^{-1}(N(z_1, z_2)ZF + M(z_1, z_2)K_f) \\ &= \begin{bmatrix} V(z_1, z_2) & \bar{W}_{ry}(z_1, z_2) \end{bmatrix} \begin{bmatrix} L(z_1, z_2) \\ S(z_1, z_2) \end{bmatrix} \begin{bmatrix} N(z_1, z_2) & M(z_1, z_2) \end{bmatrix} \begin{bmatrix} ZF \\ K_f \end{bmatrix} \\ &= \Lambda(z_1, z_2)\Gamma_f(z_1, z_2) = z_1^{n_1} z_2^{n_2} I_f. \end{aligned} \tag{3.40}$$

Equations (3.32), (3.36) and (3.40) indicate that, except for the effect of initial conditions, the residual $r(z_1, z_2)$ is explicitly related to the fault $f(z_1, z_2)$ by $r(z_1, z_2) = z_1^{n_1} z_2^{n_2} f(z_1, z_2)$. The fault can, therefore, be detected from the residual generator by monitoring the evolvement of residual values. This completes the proof. ■

For the resulting residual generator with zero initial conditions, there exist $n_1 > 0$ and $n_2 > 0$ such that

$$r(i, j) = f(i - n_1, j - n_2). \tag{3.41}$$

The fault is thus detected with possibly some delay, i.e., the method can detect a fault occurring at (i, j) only at $(i + n_1, j + n_2)$. Herein the term delay used in the 1-D behavior is extended to describe the delay or shift in either time or spatial direction. To minimize the detection delay or shift, $\Lambda(z_1, z_2)$ may be calculated by expressing the polynomial inverse of $\Gamma_f(z_1, z_2)$ as:

$$\begin{bmatrix} z_1^{-n11} z_2^{-n12} & & \\ & \ddots & \\ & & z_1^{-nf1} z_2^{-nf2} \end{bmatrix} \Lambda(z_1, z_2), \tag{3.42}$$

with $ni1$ and $ni2$ being as small as possible.

With the transfer functions of the residual generator $\bar{W}_{ry}(z_1, z_2)$, $\bar{W}_{ru}(z_1, z_2)$ and $\bar{W}_{r\hat{x}}(z_1, z_2)$ obtained, a Roesser model of the residual generator can be obtained using the realization described in the last section, leading to the state space model of the residual generator in the following form:

$$\begin{gathered} \begin{bmatrix} v^h(i+1, j) \\ v^v(i, j+1) \end{bmatrix} = \begin{bmatrix} \bar{A}_{11} & \bar{A}_{12} \\ \bar{A}_{21} & \bar{A}_{22} \end{bmatrix} \begin{bmatrix} v^h(i, j) \\ v^v(i, j) \end{bmatrix} + \begin{bmatrix} \bar{B}_1 \\ \bar{B}_2 \end{bmatrix} \begin{bmatrix} u(i, j) \\ y(i, j) \\ \hat{x}(i, j) \end{bmatrix}, \\ r(i, j) = \begin{bmatrix} \bar{C}_1 & \bar{C}_2 \end{bmatrix} \begin{bmatrix} v^h(i, j) \\ v^v(i, j) \end{bmatrix} + \bar{K} \begin{bmatrix} u(i, j) \\ y(i, j) \\ \hat{x}(i, j) \end{bmatrix}. \end{gathered} \tag{3.43}$$

The residue generator can be denoted as $\{\bar{A},\ \bar{B},\ \bar{C},\ \bar{K}\}$. From the transfer function, realization is then used to obtain The schematic of constructing the residual generator is shown in Fig. 3.3.

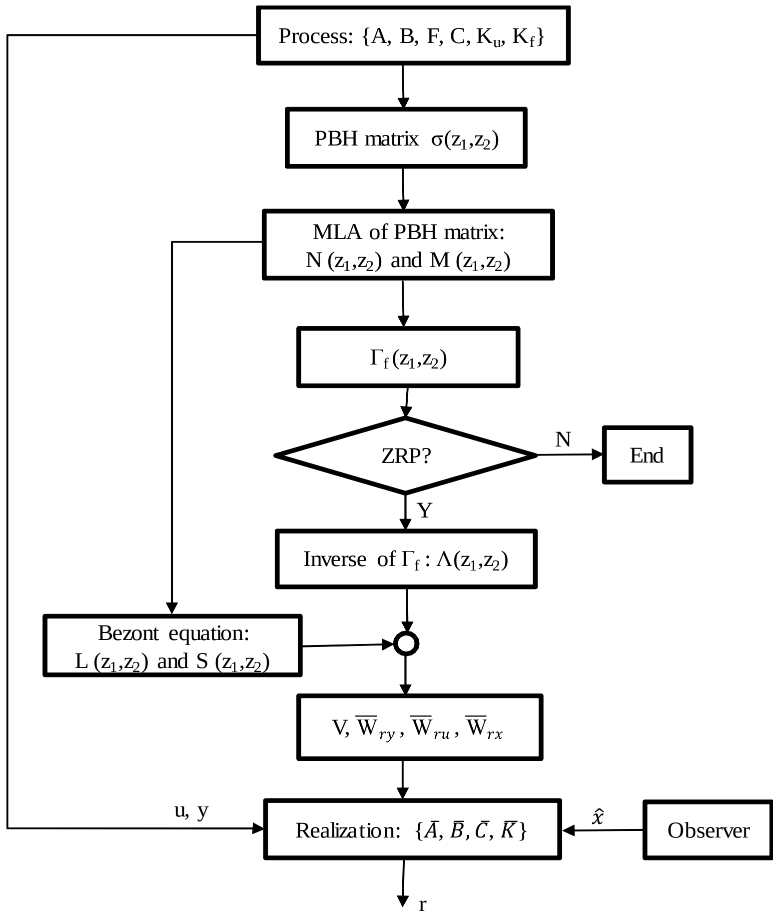


Figure 3.3: Construction of a residual generator for a 2-D Roesser system without unknown disturbances.

In monitoring an industrial process, it is essential to detect and locate faults as soon as possible. Delays, however, often exist due to monitoring mechanism or equipment. In the 2-D observer based fault detection algorithm, delays could be induced from the inverse of polynomial matrix Γ_f and realizations of observers and residue generators. In the proposed strategy, through using an efficient realization technique and expressing the inverse of $\Gamma_f(z_1, z_2)$ with (3.42), it is expected to generate efficient fault detections.

3.3.3 Systems with unknown disturbance

If there exists an unknown disturbance in the input signals, additional polynomial matrices for disturbances are necessary in constructing the fault detection strategy. When unknown disturbances affect the system dynamics, the Roesser model is expressed as:

$$\begin{aligned} \begin{bmatrix} x^h(i+1,j) \\ x^v(i,j+1) \end{bmatrix} &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x^h(i,j) \\ x^v(i,j) \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u(i,j) \\ &\quad + \begin{bmatrix} D_1 \\ D_2 \end{bmatrix} d(i,j) + \begin{bmatrix} F_1 \\ F_2 \end{bmatrix} f(i,j), \\ y(i,j) &= \begin{bmatrix} C_1 & C_2 \end{bmatrix} \begin{bmatrix} x^h(i,j) \\ x^v(i,j) \end{bmatrix} \\ &\quad + K_u u(i,j) + K_d d(i,j) + K_f f(i,j), \end{aligned} \tag{3.44}$$

where $d(i,j)$ indicates unknown disturbances and D_1, D_2 and K_d are matrices with proper dimensions.

For systems affected by unknown disturbances, determining whether the observer exists is more involved than the case without disturbances. Let

$$D = \begin{bmatrix} D_1 \\ D_2 \end{bmatrix}.$$

It is necessary to calculate the minimum left annihilator of disturbance matrix $\begin{bmatrix} ZD \\ -K_d \end{bmatrix}$ in order to determine whether an observer exists. Denote the minimum left annihilator of $\begin{bmatrix} ZD \\ -K_d \end{bmatrix}$ as $\begin{bmatrix} H_d & H_k \end{bmatrix}$. With the calculated PBH observability matrix $\sigma(z_1, z_2)$ in (3.11), a polynomial matrix can be defined:

$$\Gamma_1(z_1, z_2) = \begin{bmatrix} H_d & H_k \end{bmatrix} \sigma(z_1, z_2). \tag{3.45}$$

Existence of a dead-beat observer can be determined by examining zero right primeness of $\Gamma_1(z_1, z_2)$. If $\Gamma_1(z_1, z_2)$ is zero right prime, there exists a dead beat observer. Its transfer function can be obtained similarly to that for systems with no disturbances, as described in the last subsection. The only difference is that the polynomial pair $\begin{bmatrix} Q(z_1, z_2) & P(z_1, z_2) \end{bmatrix}$ has to be calculated satisfying a more stringent condition:

$$\begin{bmatrix} Q(z_1, z_2) & P(z_1, z_2) \end{bmatrix} \begin{bmatrix} I_x - ZA & ZD \\ C & -K_d \end{bmatrix} = \begin{bmatrix} I_x & 0 \end{bmatrix}. \tag{3.46}$$

From the transfer function, realization is then used to obtain the Roesser model of the observer.

With the state observer constructed, it is necessary to determine whether an observer based fault detection residual generator exists. By calculating the polynomial matrix pair $\begin{bmatrix} N(z_1, z_2) & M(z_1, z_2) \end{bmatrix}$ using equation (3.24), a disturbance related polynomial $\Gamma_d(z_1, z_2)$ can be determined:

$$\Gamma_d(z_1, z_2) = \begin{bmatrix} N(z_1, z_2) & M(z_1, z_2) \end{bmatrix} \begin{bmatrix} ZD \\ K_d \end{bmatrix}. \tag{3.47}$$

Denote the minimum left annihilator of $\Gamma_d(z_1, z_2)$ as $\Phi_d(z_1, z_2)$. The fault related polynomial matrix $\Gamma_f(z_1, z_2)$ is calculated using equation (3.25). The polynomial matrix is obtained:

$$\Gamma_R(z_1, z_2) = \Phi_d(z_1, z_2)\Gamma_f(z_1, z_2). \tag{3.48}$$

There exists an observer based fault detection residual generator if $\Gamma_R(z_1, z_2)$ is zero right prime. With zero right primeness of the polynomial $\Gamma_1(z_1, z_2)$ and $\Gamma_R(z_1, z_2)$ satisfied, the residual generator for fault detection can be obtained by finding a L-polynomial left inverse of $\Gamma_d(z_1, z_2)$ with equation (3.42) satisfying

$$\Lambda(z_1, z_2)\Gamma_d(z_1, z_2) = 0. \tag{3.49}$$

With the obtained $\Lambda(z_1, z_2)$, the same procedure described in the last subsection is then used to construct a residual generator. The schematic of residual generator construction is shown in Fig. 3.4.

A desirable fault detection should be insensitive to occurrence of disturbances. The proposed fault detection is robust to unknown disturbances when the stricter conditions of the observer and residual generator are satisfied. The robust property of the fault detection to unknown disturbance is further illustrated in the Example 3 of Simulation.

In contrast to the F-M model, the state variables in the Roesser model are comprised of horizontal and vertical variables. The two independent variables z_1 and z_2 corresponding to the horizontal and vertical evolvement of the variables are expressed in a compact form using the matrix Z defined in (3.2). As a result, the transfer functions, PBH matrix, Bézout equation and subsequent development of a residual generator need to be constructed to be the functions of the matrix Z. Existence of horizontal and vertical variables in the Roesser model also increases the challenge of realization. Minimal realization is necessary for implementation of the proposed technique.

3.4 Simulations

In this section, the proposed fault detection strategy is applied to three examples to examine its performance. For simplicity of calculations, the initial values of the plants, observers and residual generators were set to zero in the simulations.

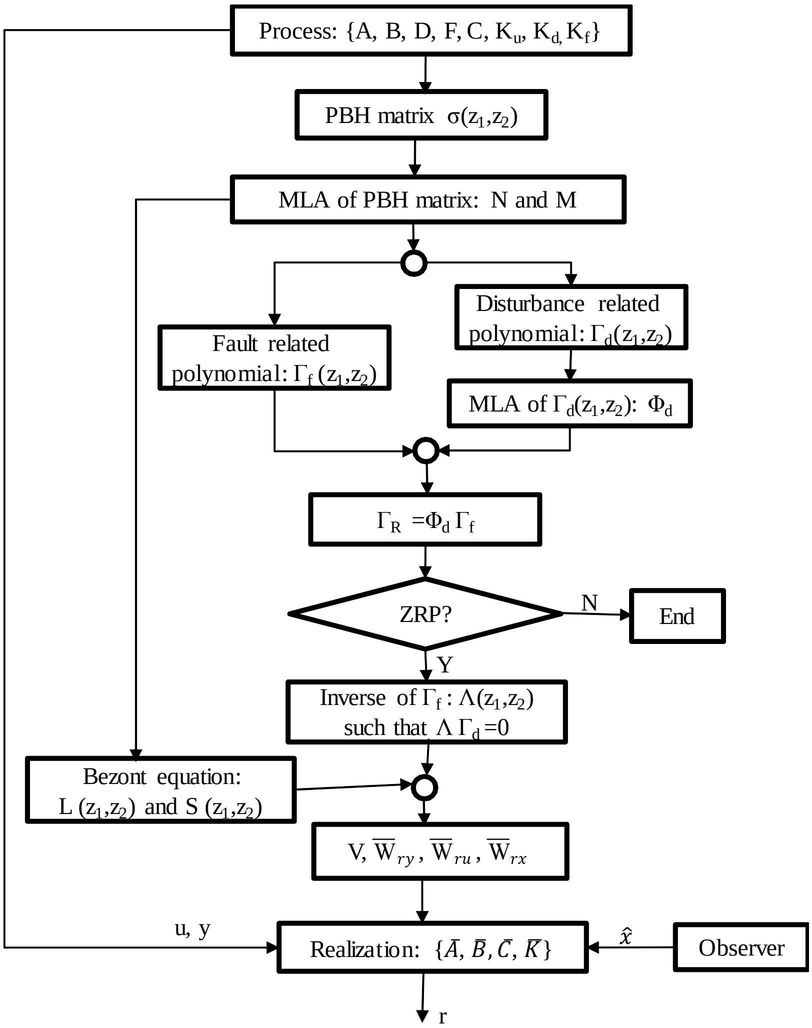


Figure 3.4: Construction of a residual generator for a 2-D Roesser system affected by unknown disturbances.

Example 1. Consider the following 2-D Roesser model with a fault signal:

$$\begin{aligned}\begin{bmatrix} x^h(i+1,j) \\ x^v(i,j+1) \end{bmatrix} &= \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x^h(i,j) \\ x^v(i,j) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(i,j) + \begin{bmatrix} -1 \\ 0 \end{bmatrix} f(i,j), \\ y(i,j) &= \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x^h(i,j) \\ x^v(i,j) \end{bmatrix}. \end{aligned} \tag{3.50}$$

From (3.50), there exists no disturbance in this example. The proposed fault detection for systems without disturbance, described in Section 3.2, can then be applied to this example. By solving the Bézout equation of the PBH observability matrix, the transfer function for the observer can be obtained:

$$\hat{W}_o = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \tag{3.51}$$

As this transfer function is very simple, the state estimate $\hat{x}$ can be directly obtained:

$$\begin{bmatrix} \hat{x}^h(i,j) \\ \hat{x}^v(i,j) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} y(i,j). \tag{3.52}$$

Based on the obtained state estimate $\hat{x}(i,j)$, the residual generator can be calculated with the state estimate $\hat{x}$, process input $u(i,j)$, process output $y(i,j)$ as the input to the residual generator and the residual $r(i,j)$ as the output. The transfer function of the residual generator for this system is:

$$\hat{W}_r = \left[\begin{array}{cc|c|c} 1-z_1 & 0 & 0 & -1+z_1 \end{array}\right]. \tag{3.53}$$

Using the realization, the state space Roesser model of the residual generator is obtained:

$$\begin{aligned}
\begin{bmatrix} v^h(i+1,j) \\ v^v(i,j+1) \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v^h(i,j) \\ v^v(i,j) \end{bmatrix} + \left[\begin{array}{cc|c|c} -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right] \begin{bmatrix} \hat{x}(i,j) \\ u(i,j) \\ y(i,j) \end{bmatrix}, \\
r(i,j) &= \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} v^h(i,j) \\ v^v(i,j) \end{bmatrix} + \left[\begin{array}{cc|c|c} 1 & 0 & 0 & -1 \end{array}\right] \begin{bmatrix} \hat{x}(i,j) \\ u(i,j) \\ y(i,j) \end{bmatrix}.
\end{aligned} \tag{3.54}$$

Simulation was carried out to evaluate the effectiveness of the fault detection strategy. In the simulation, the spacial index i is ranging from 1 to 50 and the input signal $u(i,j)$ is set to be a random variable varying from 0 to 1, as illustrated in Fig. 3.5. When a fault signal of unit step occurs at time $j = 100$, variation of the residual r can be seen in Fig. 3.6. It is observed that, at $i = 1$, r responds instantly for the step change in f and increases with time. This indicates that the occurrence of a fault can trigger changes in r and that the proposed fault detection strategy is effective for the system unaffected by unknown disturbances. However, although a unit step change in the fault signal f occurs at every spacial point for i from 1 to 50, the residual signal r displays a change only at $i = 1$ and it is constantly 0 at other spacial points for i from 2 to 50. This shows that the strategy can detect the occurrence of a fault but is not capable of diagnosing the spacial location where a fault occurs. It is also noted that there is no delay in detecting the fault for this system.

Example 2. In the last example, the system is not affected by any unknown disturbances. In some applications, however, there exist unknown disturbances. For systems affected by unknown disturbances, it requires stricter conditions in order for an observer and a fault detection residual generator to be effective. In this example, the effect of unknown disturbances on the performance of fault detection is investigated for a system that may not satisfy the stricter con-

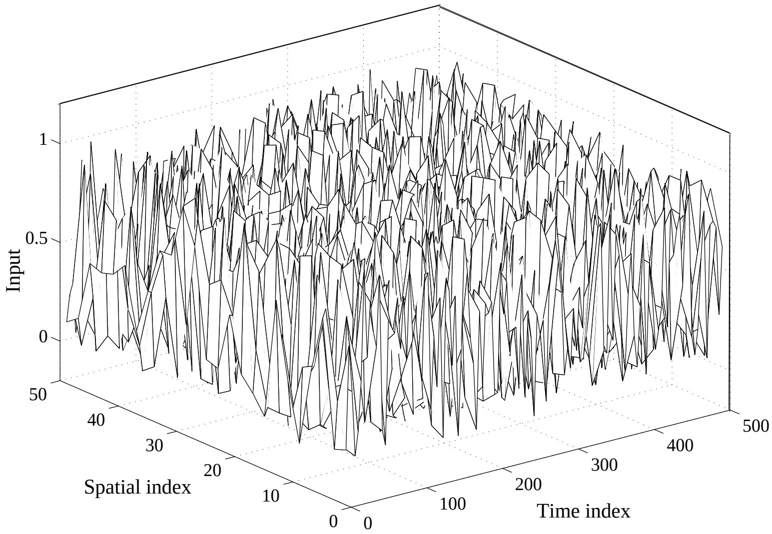


Figure 3.5: A random input signal.

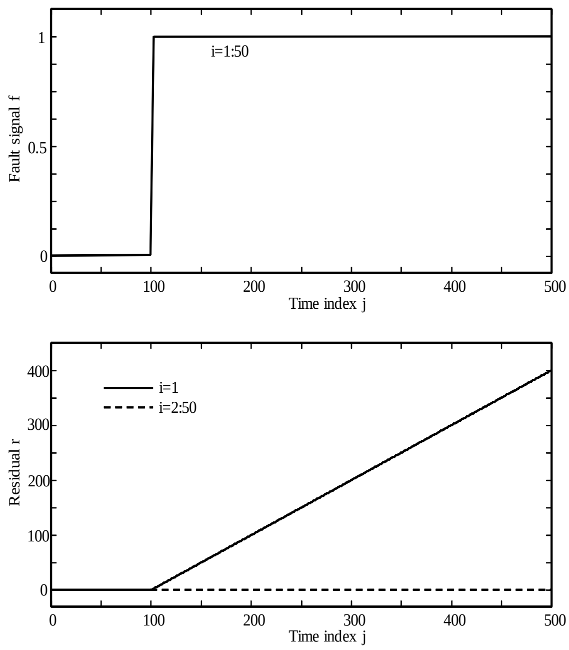


Figure 3.6: Residual response for a step change in fault signal for a system without unknown disturbances.

ditions. An example the same as the last but with an additional disturbance signal is considered:

$$\begin{aligned}
\begin{bmatrix} x^h(i+1,j) \\ x^v(i,j+1) \end{bmatrix} &= \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x^h(i,j) \\ x^v(i,j) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(i,j) + \begin{bmatrix} -1 \\ 0 \end{bmatrix} f(i,j), \\
y(i,j) &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x^h(i,j) \\ x^v(i,j) \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \end{bmatrix} d(i,j).
\end{aligned} \tag{3.55}$$

As this system is the same as that in the last example, the same observer and fault detection residual generator as in (3.52) and (3.54) can be obtained using the proposed fault detection method. In the simulation, the process input $u(i,j)$ was a random signal as in Fig. 3.5, and the unknown disturbance d was also set to be a random signal ranging from 0 to 1 as in Fig. 3.7. When a fault signal of step change occurs at time $j = 100$, variation of the residual signal is examined. Fig. 3.8 indicates that the residual signal r does not respond to occurrence of the fault signal and is constantly 0 across every spacial point. This shows that presence of the unknown input disturbance has jeopardized performance of the residual generator. Failure to detect the fault signal in this example is due to the fact that the calculated observer does not completely satisfy the conditions of a dead-beat observer.

Example 3. The performance of the fault detection strategy is further investigated for a system with unknown disturbances satisfying the stricter conditions of a dead beat observer and a residual generator. The following system affected by an unknown disturbance is considered:

$$\begin{aligned}
\begin{bmatrix} x^h(i+1,j) \\ x^v(i,j+1) \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x^h(i,j) \\ x^v(i,j) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} u(i,j) + \begin{bmatrix} 0 \\ -1 \end{bmatrix} f(i,j), \\
y(i,j) &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x^h(i,j) \\ x^v(i,j) \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \end{bmatrix} d(i,j) + \begin{bmatrix} 0 \\ -1 \end{bmatrix} f(i,j).
\end{aligned} \tag{3.56}$$

This is a system with an unknown disturbance. By calculation, it is found that Γ_1 in (3.45) and Γ_R in (3.48) are zero right prime, and a dead beat observer and a residual generator are, therefore, exist.

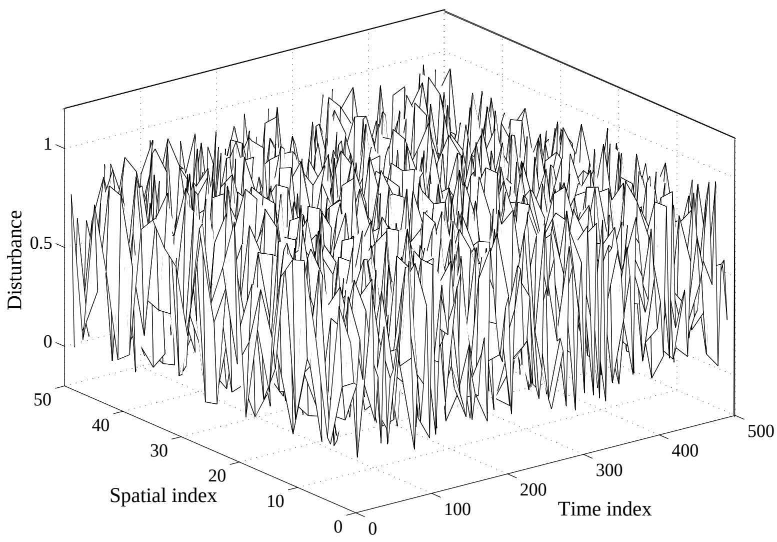


Figure 3.7: A random disturbance signal in Example 2.

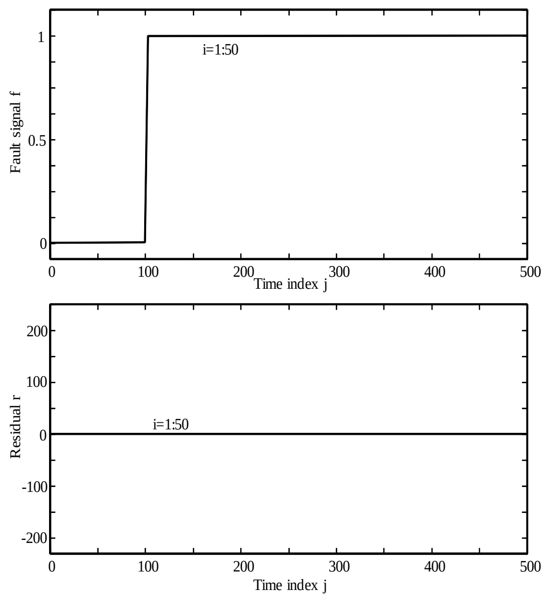


Figure 3.8: Residual response for a step change in fault signal for a system affected by unknown disturbances.

The transfer function of the observer can be obtained:

$$\hat{W}_o = \begin{bmatrix} 0 & z_1 z_2 & 0 \\ 0 & z_2 & 0 \end{bmatrix}. \tag{3.57}$$

The realization can then be used to generate the Roesser model of the observer, which takes in the following form:

$$\begin{aligned} \begin{bmatrix} w^h(i+1,j) \\ w^v(i,j+1) \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} w^h(i,j) \\ w^v(i,j) \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} u(i,j) \\ y(i,j) \end{bmatrix}, \\ \hat{x}(i,j) &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} w^h(i,j) \\ w^v(i,j) \end{bmatrix}. \end{aligned} \tag{3.58}$$

Corresponding to the state estimate $\hat{x}(i,j)$, process input $u(i,j)$, process output $y(i,j)$ as the inputs to the residual generator, the transfer function of the residual generator for this system is:

$$\bar{W}_r = \left[\begin{array}{cc|c|cc} z_1 z_2 - 1 & 0 & 0 & 0 & 0 \end{array}\right]. \tag{3.59}$$

The state space Roesser model of the residual generator can then be obtained using the realization technique:

$$\begin{aligned} \begin{bmatrix} v^h(i+1,j) \\ v^v(i,j+1) \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v^h(i,j) \\ v^v(i,j) \end{bmatrix} + \left[\begin{array}{cc|c|cc} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{array}\right] \begin{bmatrix} \hat{x}(i,j) \\ u(i,j) \\ y(i,j) \end{bmatrix}, \\ r(i,j) &= \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} v^h(i,j) \\ v^v(i,j) \end{bmatrix} + \left[\begin{array}{cc|c|cc} -1 & 0 & 0 & 0 & 0 \end{array}\right] \begin{bmatrix} \hat{x}(i,j) \\ u(i,j) \\ y(i,j) \end{bmatrix}. \end{aligned} \tag{3.60}$$

Simulation was run for both the cases without disturbance and with a disturbance. In the simulation, the process input u was set to be a random signal ranging from 0 to 1. Fig. 3.9 and 3.10 illustrates the residual response to a step change in the fault signal f when there is no

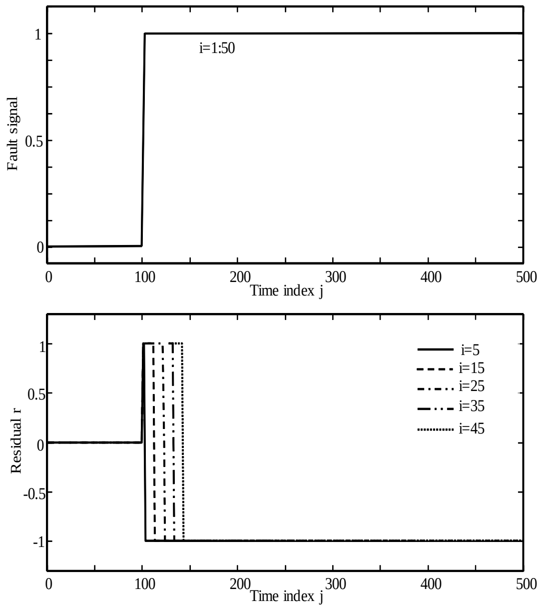


Figure 3.9: Fault residual response for a step change in fault signal for a system without unknown disturbances.

disturbance, i.e., $d = 0$. It is observed that the residual r responds quickly to occurrence of the fault and displays some delays at different spacial locations. The residual varies in the range between -1 and 1.

Effect of unknown disturbances was examined by introducing a step disturbance signal at time $j = 300$, as in Fig 3.11. The evolvement of the residual is displayed in Fig. 3.12 and 3.13. It is noted that, for this example, the residual responds to the fault signal despite the disturbance and settles down at different values at different spacial points. Introduction of disturbance at time $j = 300$ does not affect the residual response. The proposed strategy is robust to disturbances and capable of detecting faults for systems without disturbance as well as systems with disturbances.

The performance of the fault detection is also investigated for the case when the fault signal changes from occurrence to vanishing. Such a change of the fault signal can be represented by a pulse function. Fig. 3.14 and 3.15 show the residual response to a fault signal of the pulse function. It is observed that the residual increases and settles at constant values when the fault occurs at time $j = 100$, and the residual starts to decreases and settles at zero when the fault vanishes at time $j = 300$. The result indicates that the residual signal is sensitive to fault

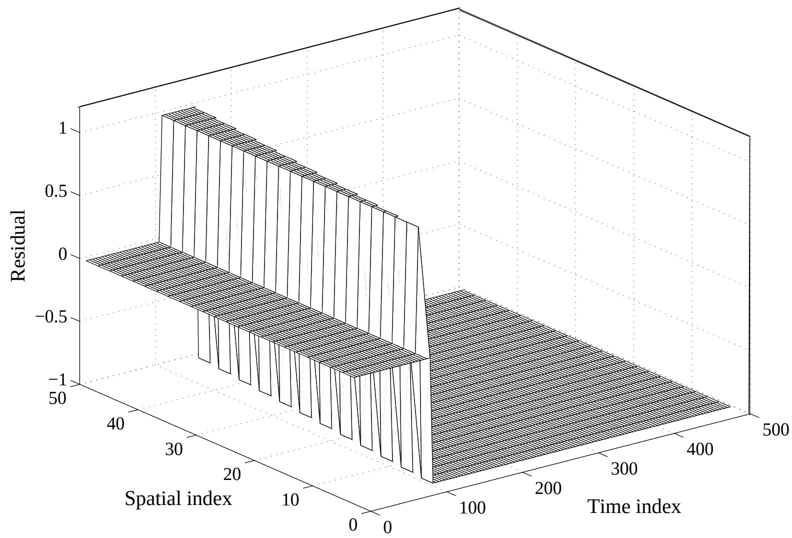


Figure 3.10: 3-D plot of residual response for a step change in fault signal for a system without unknown disturbances.

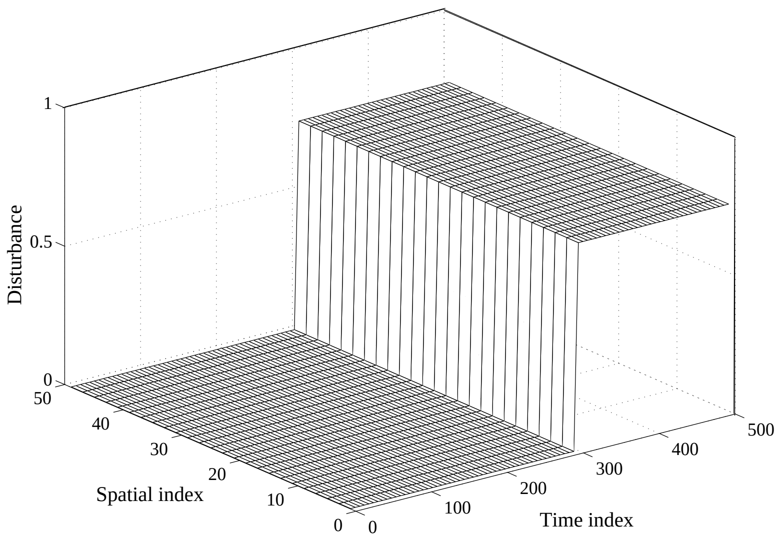


Figure 3.11: A disturbance signal in Example 3.

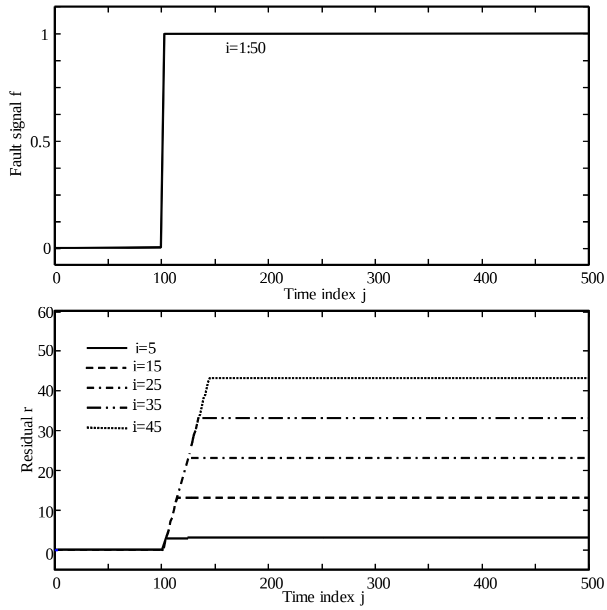


Figure 3.12: Residual response for a step change in fault signal for a system affected by unknown disturbances.

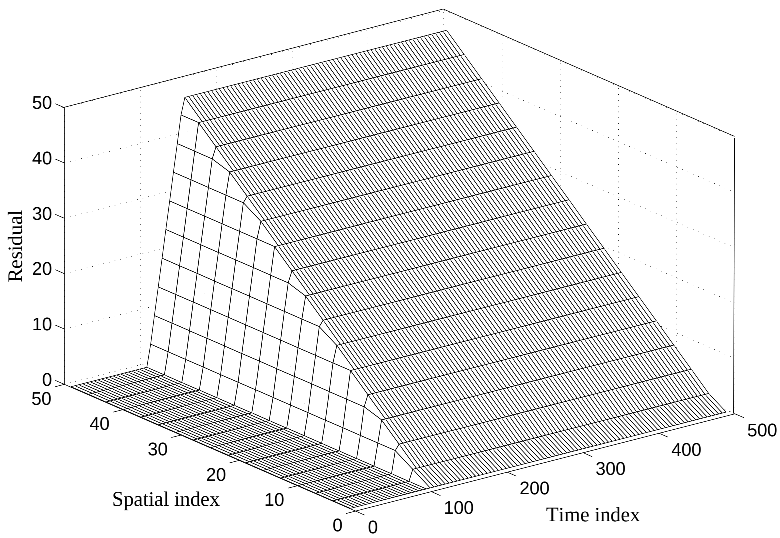


Figure 3.13: 3-D plot of residual response for a step change in fault signal for a system affected by unknown disturbances.

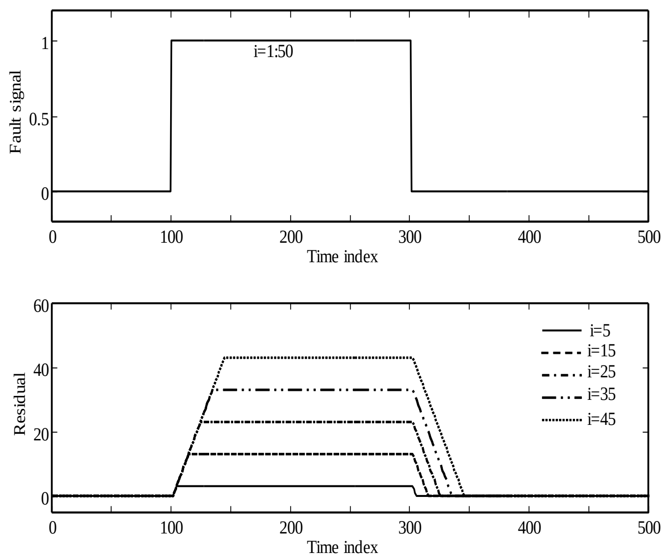


Figure 3.14: Residual response for a pulse change in fault signal for a system affected by unknown disturbances.

occurrence and vanishing and can, therefore, be used to detect the faults with reliability.

3.5 Summary

Fault detection for 2-D systems is a challenging topic and only limited research results have been reported. In this chapter, an observer based fault detection method is developed for systems modeled by the Roesser models. The proposed method is comprised of construction of a dead beat observer and a residual generator. The PBH observability matrix of a system is calculated, and some system specific polynomial matrices are then formed and their zero right primeness is used to determine the existence of the observer and the residual generator. The systems affected by unknown disturbances require more stringent conditions. Realization of the Roesser state space models from transfer functions is an integral part of the fault detection method. An realization technique generating the state space models of minimal order is integrated and leads to an efficient fault detection strategy by reducing the computation cost. Simulation results show that the algorithm is capable of detecting faults for systems described by the Roesser model. Disturbance may jeopardize the performance of the fault detection strategy. Fault detection

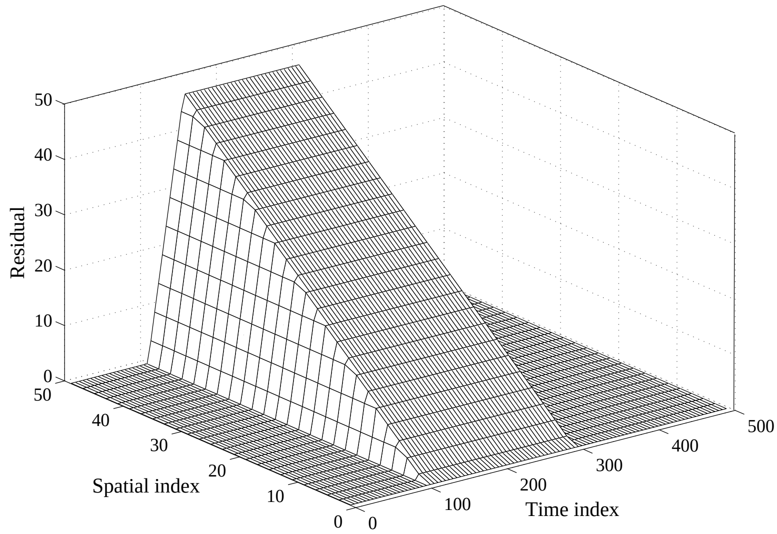


Figure 3.15: 3-D plot of fault residual response for a pulse change in fault for a system affected by unknown disturbances.

of 2-D processes represented by the Roesser Model represents a necessary extension of the existing fault detection methods.

The proposed algorithm is validated using numerical examples. Some challenges, however, still exist in implementing the technique into real world applications. One of them lies in the difficulties of model identification based on industrial data. Future work should include implementing the developed fault detection into industrial applications as well as 2-D model identification based on industrial data.

Chapter 4

Kalman Filter Based FDI for F-M Model in Two Dimensional Systems

4.1 Introduction

Fault detection and diagnosis are essential in ensuring safe operations by providing timely diagnostic information to plant operators in process industry. Extensive research has been carried out on fault detection and diagnosis ranging from analytical methods to artificial intelligence and statistical approaches. The analytical schemes for fault detection and diagnosis are basically signal processing techniques using state estimation, parameter estimation and adaptive filtering (Venkatasubramanian et al., 2003). The key component in these schemes is to design diagnostic observers with satisfactory decoupling properties for residual generation of dynamic systems. For systems where variables display noisy fluctuations with known statistical parameters, fault diagnosis problem entails monitoring the innovation process or the prediction errors (Basseville, 1988; Willsky and Jones, 1976). A Kalman filter, designed on the basis of the system model in its normal operating mode, provides a state estimator with minimum estimation error and has been used for the fault detection and isolation purposes (Fathi et al., 1993; Chang and Hwang, 1998). The well-developed fault detection and diagnosis techniques have been predominantly focused on 1-D system.

Research and development on FDI of 2-D systems have been limited, owing in part to complexity of 2-D models (Wu et al., 2011, 2012; Maleki et al., 2013). The dead-beat observer has recently been applied to fault detection and isolation of 2-D systems by constructing a residual generator (Bisiaco and Valcher, 2004, 2006). The dead-beat observer based FDI was shown to be effective for 2-D systems but strict conditions are required in order for an observer and a residual generator to exist. These strict conditions may not be satisfied for some systems.

A Kalman filter, which minimizes the estimation error variance, has become one of the most popular state estimation approaches. The Kalman filter based fault detection methods have also been well accepted for 1-D cases. Since 1970s, extensive research has been involved with the attempt to introduce 2-D Kalman filters (Shanks et al., 1972; Woods and Radewan, 1977; Katayama and Kosaka, 1979; Sebek, 1992). A straightforward extension of 1-D Kalman filter techniques would result in a number of state variables proportional to N for the filtering of an $N \times N$ digital image (Sheng and Zou, 2007; Kwan and Lewis, 1999; Katayama and Kosaka, 1979). In a 2-D case, the enormous quantity of the data calls for an efficient recursion processor. In parallel with the active research on development of efficient 2-D Kalman filters, it is of great interest to explore the applications of 2-D Kalman filters for fault detection of 2-D systems.

In this chapter, a Kalman filter based fault detection method is developed for 2-D systems described by F-M models. The state estimate minimizing the estimation error variance is recursively calculated using a 2-D Kalman filter. The obtained state estimate leads to generation of a residual through the innovation process. The generated residual is of zero mean Gaussian noise when there exists no fault, and is no longer zero mean Gaussian noise when there are faults. The residual directly reflects the fault information and is therefore used to determine whether a fault occurs. The model of the residual over a moving evaluation window is formulated describing the relation of the residual with the fault. Simulations were carried out on two examples to investigate the performance of the proposed fault detection method. In both examples, the generated residual responds quickly to occurrences of faults and the proposed method is shown to be effective in fault detection of 2-D systems. The Kalman filter based fault detection does not require to satisfy the strict conditions and can, therefore, be used to a wider range of 2-D systems.

4.2 2-D Kalman filter based fault detection

In the Kalman filter based fault detection, a residual signal is generated using the state estimation from a Kalman filter. Due to the noisy characteristics of the systems, it is necessary to examine the residual over a 2-D evaluation window and then detects with its statistical properties.

4.2.1 Residual generation

A recursive Kalman filter is used to derive a state estimate and the residual is then calculated based on the obtained state estimate and measured output. The structure of the 2-D Kalman filter based residual generator is illustrated in Fig. 4.1.

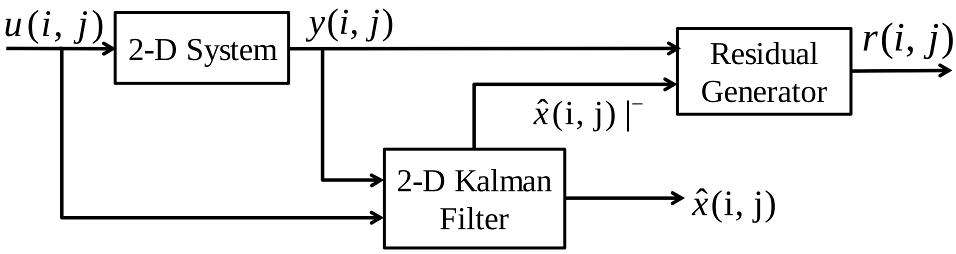


Figure 4.1: Structure of a Kalman filter based residual generator.

Consider a 2-D F-M model with a fault signal and process noises:

$$\begin{aligned} x(i+1,j+1) = \quad & A_1x(i+1,j) + A_2x(i,j+1) + B_1u(i+1,j) + B_2u(i,j+1) \\ & +F_1f(i+1,j) + F_2f(i,j+1) + H_1\eta(i+1,j) + H_2\eta(i,j+1) \\ y(i,j) = \quad & Cx(i,j) + K_ff(i,j) + \nu(i,j) \end{aligned} \tag{4.1}$$

with the boundary conditions:

$$x(i,0) = x_{i0}, x(0,j) = x_{0j},$$

where i and j indicate the horizontal (or spacial) and vertical (or time) indices, $x(i,j) \in \mathbb{R}^{n_x}$, $u(i,j) \in \mathbb{R}^{n_u}$, $f(i,j) \in \mathbb{R}^{n_f}$and $y(i,j) \in \mathbb{R}^{n_y}$ are state, input, fault and output vectors, respectively, $\eta(i,j)$ indicates the process noise, $\nu(i,j)$ is measurement noise, A_1, A_2, B_1, B_2, H_1, H_2, F_1, F_2, C, K_f are known matrices of appropriate dimensions. It is assumed that the variables $\eta(i,j) \in \mathbb{R}^{n_\eta}$, and $\nu(i,j) \in \mathbb{R}^{n_\nu}$ be white Gaussian noises with zero mean value and given variances:

$$
\begin{aligned}
&E\left[\eta(i,j)\right] = 0, \\
&E\left[\nu(i,j)\right] = 0, \\
&E\left[\eta(i,j)\eta^T(k,l)\right] = \begin{cases} Q(i,j), & \text{when } i = k \text{ and } j = l, \\ 0, & \text{elsewhere }, \end{cases} \\
&E\left[\nu(i,j)\nu^T(k,l)\right] = \begin{cases} R(i,j), & \text{when } i = k \text{ and } j = l, \\ 0, & \text{elsewhere }, \end{cases}
\end{aligned}
\tag{4.2}
$$

where Q, and R are the covariance matrices of $\eta(i,j)$ and $\nu(i,j)$, respectively.

Let $\widehat{x}(i,j)$ be the estimate of $x(i,j)$ and $\widetilde{x}(i,j)$ be the estimation error, i.e., $\widetilde{x}(i,j) = x(i,j) - \widehat{x}(i,j)$. A Kalman filter is designed to update state estimate $\widehat{x}(i,j)$ such that the variance of estimation error $\widetilde{x}(i,j)$ be minimized. The variance of estimate error $\widetilde{x}(i,j)$ can be written as:

$$
P(i,j) = E\left[\widetilde{x}(i,j)\widetilde{x}(i,j)^T\right] \tag{4.3}
$$

With the given initial condition, the boundary values of $P(i,j)$ are obtained:

$$
\begin{aligned}
P(i,0) &= E\left[(x_{i0} - \widehat{x}(i,0))(x_{i0} - \widehat{x}(i,0))^T\right] \\
P(0,j) &= E\left[(x_{0j} - \widehat{x}(0,j))(x_{i0} - \widehat{x}(0,j))^T\right]
\end{aligned}
\tag{4.4}
$$

The proposed fault detection is to generate a residual signal such that it reflects occurrence of faults. For the residual generation, the state estimate $\widehat{x}(i,j)$ is to be updated recursively. The Kalman filter proposed in (Zou et al., 2004) is used to update the state estimate in this work. Although there are other structures of Kalman filter in literature (Yang et al., 2009; Kwan and Lewis, 1999; Woods, 1979; Woods and Ingle, 1981; Azimi-Sadjadi and Wong, 1987; Azimi-

Sadjadi and Bannour, 1991), the genuine recursive approach is still not fully accessible. The recursive scheme for optimal state estimate can be expressed:

$$\widehat{x}(i,0) = \widehat{x}_{i0}, \qquad \widehat{x}(0,j) = \widehat{x}_{0j} \tag{4.5}$$

$$\begin{aligned} \widehat{x}(i,j)|^{-} &= A_1\widehat{x}(i,j-1) + A_2\widehat{x}(i-1,j) \\ &+B_1u(i,j-1) + B_2u(i-1,j) \end{aligned} \tag{4.6}$$

$$\widehat{x}(i,j) = \widehat{x}(i,j)|^{-} + L(i,j)[y(i,j) - C\widehat{x}(i,j)|^{-}] \tag{4.7}$$

where $\widehat{x}(i,j)|^{-}$ indicates the state estimate based on measurement up to $(i-1,j)$ and $(i,j-1)$, while $\widehat{x}(i,j)$ is the state estimate based on measurement up to (i,j) and $L(i,j)$ denotes the Kalman filter gain. Corresponding to $\widehat{x}(i,j)|^{-}$ and $\widehat{x}(i,j)$, the associated state estimation error variance matrices are:

$$P(i,j)|^{-} = E\left[(x(i,j) - \widehat{x}|^{-}(i,j))(x(i,j) - \widehat{x}|^{-}(i,j))^T\right] \tag{4.8}$$

$$P(i,j) = E\left[(x(i,j) - \widehat{x}(i,j))(x(i,j) - \widehat{x}(i,j))^T\right] \tag{4.9}$$

In updating the state estimate $\widehat{x}(i,j)$ in (4.7), the Kalman filter gain $L(i,j)$ needs to be calculated by recursively updating the state estimate error variances $P(i,j)|^{-}$ and $P(i,j)$:

$$\begin{aligned} P(i,j)|^{-} &= A_1P(i,j-1)A_1^T + A_2P(i-1,j)A_2^T \\ &\quad +A_1\text{cov}[\tilde{x}(i,j-1),\tilde{x}(i-1,j)]A_2^T + A_2\text{cov}[\tilde{x}(i-1,j),\tilde{x}(i,j-1)]A_1^T \\ &\quad +H_1Q(i,j-1)H_1^T + H_2Q(i-1,j)H_2^T \end{aligned} \tag{4.10}$$

$$L(i,j) = P(i,j)|^{-}C^T\left[CP(i,j)|^{-}C^T + R(i,j)\right]^{-1} \tag{4.11}$$

$$P(i,j) = (I - L(i,j)C)P(i,j)|^{-} \tag{4.12}$$

In the above equation, the covariance matrices $\text{cov}[\tilde{x}(i,j-1),\tilde{x}(i-1,j)]$ can be assessed iteratively from those of the three neighboring points:

$$
\begin{aligned}
&\text{cov}[\tilde{x}(i,j-1),\tilde{x}(i-1,j)] = (I-L(i,j-1)C) \\
&\quad [A_1\text{cov}[\tilde{x}(i,j-2),\tilde{x}(i-1,j-1)]A_1^T + A_1\text{cov}[\tilde{x}(i,j-2),\tilde{x}(i-2,j)]A_2^T \\
&\quad +A_2\text{cov}[\tilde{x}(i-1,j-1),\tilde{x}(i-1,j-1)]A_1^T \\
&\quad +A_2\text{cov}[\tilde{x}(i-1,j-1),\tilde{x}(i-2,j)]A_2^T \\
&\quad +H_2Q(i-1,j-1)H_1^T](I-L(i-1,j)C)^T
\end{aligned}
\tag{4.13}
$$

With the state estimated calculated, the residual $r(i,j)$ is obtained from the innovation process

$$r(i,j) = y(i,j) - C\hat{x}(i,j)|^- \tag{4.14}$$

It follows from the above equations that

$$
\begin{aligned}
\tilde{x}(i,j) \;&=\; (I-L(i,j)C)(A_1\tilde{x}(i,j-1) + A_2\tilde{x}(i-1,j) \\
&\quad +F_1f(i,j-1) + F_2f(i-1,j) + H_1\eta(i,j-1) + H_2\eta(i-1,j)) \\
&\quad -L(i,j)K_ff(i,j) - L(i,j)\nu(i,j) && (4.15) \\
r(i,j) \;&=\; C(A_1\tilde{x}(i,j-1) + A_2\tilde{x}(i-1,j) + F_1f(i,j-1) + F_2f(i-1,j) \\
&\quad +H_1\eta(i,j-1) + H_2\eta(i-1,j)) + K_ff(i,j) + \nu(i,j) && (4.16)
\end{aligned}
$$

Equation (4.15) and (4.16) describe how the state estimate error and residual signal evolve along two directions. For an effective Kalman filter, the state estimate error should become a zero mean white Gaussian noise after a certain time. Consequently, if in fault free mode, i.e., $f(i,j) = 0$, the residual output $r(i,j)$ should be a zero mean white Gaussian process. When a fault occurs, i.e., $f(i,j) \neq 0$, $r(i,j)$ is no longer a zero mean white noise. The fault can be detected by evaluating the signal $r(i,j)$.

Residual generation is critical for the Kalman filter based fault detection and its on-line implementation can be performed with the following steps:

Step 1. Assuming the boundary values of a 2-D system: $x(i,0) = x_{i0}$, $x(0,j) = x_{0j}$, the state estimates at the boundaries are set up as: $\hat{x}(i,0) = x_{i0}$, $\hat{x}(0,j) = x_{0j}$. Calculate the state estimate error matrices at the boundaries $P(i,0)$ and $P(0,j)$ using (4.4).

Step 2. Calculate the covariance matrices at the boundaries

$cov[\tilde{x}(i,0), \tilde{x}(i-1,1)]$ and $cov[\tilde{x}(0,j), \tilde{x}(1,j-1)]$, $i = 1.2, \cdots$, and $j = 1, 2, \cdots$.

Step 3. Compute the covariance matrixes

$cov[\tilde{x}(i-1,j), \tilde{x}(i,j-1)]$ using Equation (4.13).

Step 4. Evaluate the matrices

$P(i,j)|^-$, $L(i,j)$ and $P(i,j)$ using Equation (4.10)-(4.12).

Step 5. Compute the state estimate $\hat{x}(i,j)|^-$ and $\hat{x}(i,j)$ using Equation (4.6) and (4.7).

Step 6. Calculate the residual signal $r(i,j)$ using (4.14).

Step 7. Evaluate the obtained $r(i,j)$ and check if it is a zero mean white noise.

Step 8. Increase i or j, and go to Step 3.

4.2.2 Model of Residual over an Evaluation Window

The residual signal $r(i,j)$ obtained from (4.14) carries the information of possible faults and can, therefore, be used to determine if a fault has occurred. Since the residual signal is corrupted by stochastic noises, it is necessary to evaluate the residual over an evaluation window. In 1-D cases, an evaluation window with a length t is defined by a time range from a specified time $j-t$ to current time j. For 2-D systems, however, an evaluation window should be defined by both horizontal and vertical directions to reflect the two dimensional evolvement of variables. In this chapter, the evaluation window is defined by a rectangular plane with the horizontal range from $i-s$ to i and the vertical range from $j-t$ to j, as illustrated in Fig. 4.2. The evaluation window moves horizontally or vertically as new measurement becomes available. In this section, the mean of the residual r and its relation with fault signals over the evaluation window are examined.

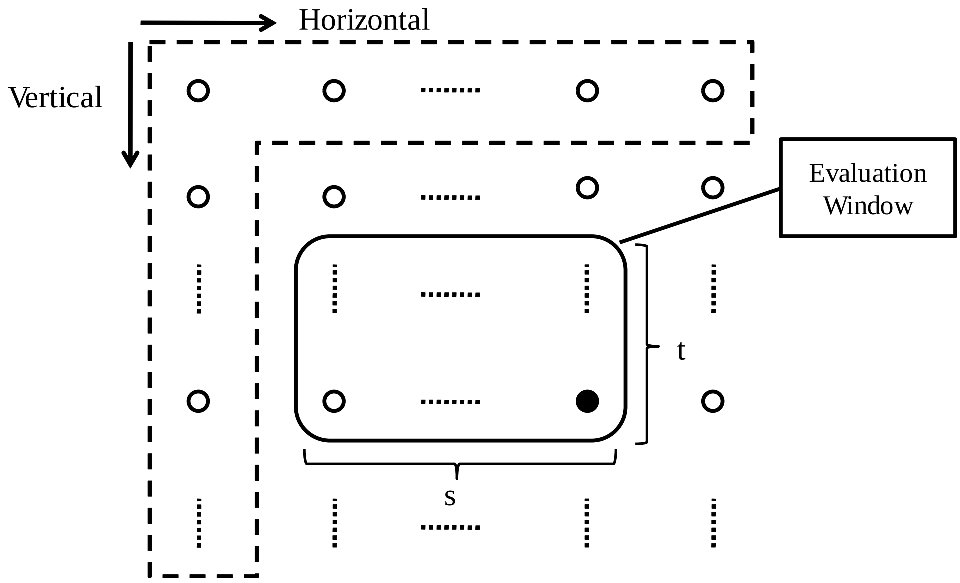


Figure 4.2: Evaluation window for 2-D residual evaluation.

Let $e(i, j)$ denote the mean of the state estimation error $\tilde{x}(i, j)$. For simplicity of description, it is assumed that the Kalman filter gain $L(i, j)$ in (4.11) converges to a constant matrix L. From (4.15), it follows that

$$\begin{aligned} e(i,j) =& \bar{A}_1 e(i, j-1) + \bar{A}_2 e(i-1, j) \\ &+ \bar{F}_1 f(i, j-1) + \bar{F}_2 f(i-1, j) - LK_f f(i, j), \end{aligned} \tag{4.17}$$

where

$$\begin{aligned} \bar{A}_1 &= (I - LC)A_1, \\ \bar{A}_2 &= (I - LC)A_2, \\ \bar{F}_1 &= (I - LC)F_1, \\ \bar{F}_2 &= (I - LC)F_2. \end{aligned}$$

In determining how the residual value is related with the fault within the evaluation window, it is necessary to formulate the relation of the estimation error with the fault. Propagation of the estimation error within the evaluation window can be formulated row by row. The estimation error and the fault in the current row are denoted by:

$$\mathbf{e}(i,i-s+1;j)=\begin{bmatrix} e(i,j)\\ e(i-1,j)\\ \vdots\\ e(i-s+1,j)\end{bmatrix},\quad \mathbf{f}(i,i-s+1;j)=\begin{bmatrix} f(i,j)\\ f(i-1,j)\\ \vdots\\ f(i-s+1,j)\end{bmatrix},$$

In describing the estimation error and residual in the range of evaluation window, a series of matrices containing zeros need to be defined. For clarity and simplicity of expression in the following description, x, y and f, when appear as subscripts, denote the dimensions of zero sub-matrices, e.g., $O_{x\times y}$ denoting a zero matrix with dimension of $n_x \times n_y$. For determining evolvement of the estimation error from one row to the next, the following matrices are defined:

$$\phi_s=\begin{bmatrix} \bar{A}_2{}^{s}\\ \bar{A}_2{}^{(s-1)}\\ \vdots\\ \bar{A}_2\end{bmatrix},$$

$$V_s=\begin{bmatrix} \bar{A}_1 & \bar{A}_2\bar{A}_1 & \cdots & \bar{A}_2{}^{(s-1)}\bar{A}_1\\ 0_{x\times x} & \bar{A}_1 & \cdots & \bar{A}_2{}^{(s-2)}\bar{A}_1\\ & & \ddots & \\ 0_{x\times x} & 0_{x\times x} & \cdots & \bar{A}_1\end{bmatrix},$$

$$\alpha_s=\begin{bmatrix} LK_f & \bar{F}_2 & 0_{x\times f} & \cdots & 0_{x\times f} & 0_{x\times f}\\ 0_{x\times f} & LK_f & \bar{F}_2 & \cdots & 0_{x\times f} & 0_{x\times f}\\ & & & \ddots & & \\ 0_{x\times f} & 0_{x\times f} & 0_{x\times f} & \cdots & LK_f & \bar{F}_2\end{bmatrix},$$

$$\beta_s=\begin{bmatrix} \bar{F}_1 & 0_{x\times f} & 0_{x\times f} & \cdots & 0_{x\times f} & 0_{x\times f}\\ 0_{x\times f} & \bar{F}_1 & 0_{x\times f} & \cdots & 0_{x\times f} & 0_{x\times f}\\ & & & \ddots & & \\ 0_{x\times f} & 0_{x\times f} & 0_{x\times f} & \cdots & \bar{F}_1 & 0_{x\times f}\end{bmatrix},$$

where 0s in the above matrices indicate the zero matrices with appropriate dimension. The estimation error at the current row is thus obtained:

$$\begin{aligned}\mathbf{e}(i, i-s+1; j) = V_s \mathbf{e}(i, i-s+1; j-1) + \phi_s e(i-s, j) \\ +\alpha_s \mathbf{f}(i, i-s; j) + \beta_s \mathbf{f}(i, i-s; j-1).\end{aligned} \tag{4.18}$$

Note that the estimation error at the current row is affected by the estimation error at the last row as well as the fault at both current and last rows. Applying Equation (4.18) iteratively yields:

$$\begin{aligned}\mathbf{e}(i, i-s+1; j) = V_s^t \mathbf{e}(i, i-s+1; j-t) \\ +\omega_{s,t} \mathbf{e}(i-s; j, j-t+1) + \alpha_s \mathbf{f}(i, i-s; j) \\ +\sigma_{s,t} \begin{bmatrix} \mathbf{f}(i, i-s; j-1) \\ \cdots \\ \mathbf{f}(i, i-s; j-t) \end{bmatrix},\end{aligned} \tag{4.19}$$

where

$$\omega_{s,t} = \begin{bmatrix} \phi_s, & V_s \phi_s, & , \cdots, & V_s^{t-1} \phi_s \end{bmatrix}, \tag{4.20}$$

$$\sigma_{s,t} = \begin{bmatrix} V_s \alpha_s + \beta_s, & , \cdots, & V_s^t \alpha_s + V_s^{t-1} \beta_s \end{bmatrix}. \tag{4.21}$$

Equation (4.19) indicates that the estimation error at the current row is a function of the estimation error at the boundaries of the evaluation window and the fault within the evaluation window. For an effective Kalman filter, the mean of the estimation error converges to zero if there is no fault. The estimation error at the boundaries of evaluation window in (4.19) is caused by the faults having occurred before the evaluation window. The mean of estimation error at the current row is, therefore, generated solely by faults within or before the evaluation window.

From the obtained relation between estimation error and fault, effects of faults on the residual signal within the evaluation window can then be obtained. Let $\bar{r}(i, j)$ denote the mean of residual $r(i,j)$, and from (4.16), it can be expressed as:

$$\begin{aligned}\bar{r}(i,j) &= \bar{C}_s\mathbf{e}(i,i-s+1;j-1) + CA_2^s e(i-s,j)\\ &+CF_1 f(i,j-1) + CF_2 f(i-1,j) + K_f f(i,j),\end{aligned} \tag{4.22}$$

where

$$\bar{C}_s = \begin{bmatrix} CA_1 & CA_2A_1 & \cdots & CA_2^{s-1}A_1 \end{bmatrix}.$$

For residual evaluation, it is necessary to examine the residual value and its relation with faults in the scope of evaluation window. Corresponding to the effect of faults on the mean of residual $\bar{r}(i,j)$, the following two matrices are defined:

$$\zeta_s = \begin{bmatrix} K_f & CF_2 & 0_{y\times f} & \cdots & 0_{y\times f} & 0_{y\times f} \\ 0_{y\times f} & K_f & CF_2 & \cdots & 0_{y\times f} & 0_{y\times f} \\ & & & \ddots & & \\ 0_{y\times f} & 0_{y\times f} & 0_{y\times f} & \cdots & K_f & CF_2 \end{bmatrix}$$

$$\bar{\alpha}_s = \bar{C}_s\alpha_s + \begin{bmatrix} CF_1 & 0_{y\times f} & \cdots & 0_{y\times f} & 0_{y\times f} \\ 0_{y\times f} & CF_1 & \cdots & 0_{y\times f} & 0_{y\times f} \\ & & \ddots & & \\ 0_{y\times f} & 0_{y\times f} & \cdots & CF_1 & 0_{y\times f} \end{bmatrix}.$$

The matrices ϕ_s, V_s, α_s and β_s are defined to describe evolvement of the mean of estimation error in (4.19). These matrices are used to specify how the estimation error at the boundaries of the evaluation window and the faults within the evaluation window affect the estimation error at the current time row j. To describe how the estimation error at the boundaries of the evaluation window and the faults in the evaluation window affect the estimation error at the time row $j-k$, the matrices ϕ_{s-k}, V_{s-k}, α_{s-k}, β_{s-k}, which are of the same dimension as that of ϕ_s, V_s, α_s and

β_s, can be defined for $k = 0, \cdots, s-1$, correspondingly:

$$\phi_{s-k} = \begin{bmatrix} 0_{k(x)\times x} \\ \bar{A}_2^{\,s-k-1} \\ \bar{A}_2^{\,(s-k-2)} \\ \vdots \\ \bar{A}_2 \end{bmatrix},$$

$$V_{s-k} = \begin{bmatrix} 0_{k(x)\times k(x)} & 0_{k(x)\times x} & \cdots & 0_{k(x)\times x} \\ 0_{x\times k(x)} & \bar{A}_1 & \cdots & \bar{A}_2^{\,(s-k-1)}\bar{A}_1 \\ 0_{x\times k(x)} & 0_{x\times x} & \cdots & \bar{A}_2^{\,(s-k-2)}\bar{A}_1 \\ & & \ddots & \\ 0_{x\times k(x)} & 0_{x\times x} & \cdots & \bar{A}_1 \end{bmatrix},$$

$$\alpha_{s-k} = \begin{bmatrix} 0_{k(x)\times k(f)} & 0_{k(x)\times f} & 0_{k(x)\times f} & \cdots & 0_{k(x)\times f} \\ 0_{x\times k(f)} & LK_f & \bar{F}_2 & \cdots & 0_{x\times f} \\ 0_{x\times k(f)} & 0_{x\times f} & LK_f & \cdots & 0_{x\times f} \\ & & & \ddots & \\ 0_{x\times k(f)} & 0_{x\times f} & 0_{x\times f} & \cdots & \bar{F}_2 \end{bmatrix},$$

$$\beta_{s-k} = \begin{bmatrix} 0_{k(x)\times k(f)} & 0_{k(x)\times f} & \cdots & 0_{k(x)\times f} & 0_{k(x)\times f} \\ 0_{x\times k(f)} & \bar{F}_1 & \cdots & 0_{x\times f} & 0_{x\times f} \\ & & & \ddots & \\ 0_{x\times k(f)} & 0_{x\times f} & \cdots & \bar{F}_1 & 0_{x\times f} \end{bmatrix}.$$

Based on the definition of V_k, ϕ_k, α_k β_k for $k = 1, \cdots, s$, similar to $\omega_{s,t}$ and $\sigma_{s,t}$ in (4.20) and (4.21), $\omega_{k,l}$ and $\sigma_{k,l}$ are defined for $k = 1, \cdots, s$, $l = 1, \cdots, t$:

$$\omega_{k,l} = \Big[\phi_k, \quad V_s\phi_k, \quad ,\cdots, \quad V_k^{l-1}\phi_k\Big] \tag{4.23}$$

$$\sigma_{k,l} = \Big[V_k\alpha_k + \beta_k, \quad ,\cdots, \quad V_k^{l}\alpha_k + V_k^{l-1}\beta_k\Big] \tag{4.24}$$

$\bar{C}_s$, ζ_s and $\bar{\alpha}_s$ are defined above to describe the effects of the estimation error and faults on the current residual $\bar{r}(i,j)$ in (4.22). To represent the residual $\bar{r}$ at different points within the evaluation window, $\bar{C}_{s-k}$, ζ_{s-k} and $\bar{\alpha}_{s-k}$, which have the same dimension as that of $\bar{C}_s$, ζ_s and $\bar{\alpha}_s$, can be defined for $k = 0, \cdots, s-1$:

$$
\begin{aligned}
\bar{C}_{s-k} &= \begin{bmatrix} 0_{y\times k(x)} & CA_1 & CA_2A_1 & \cdots & CA_2^{s-k-1}A_1 \end{bmatrix}, \\
\zeta_{s-k} &= \begin{bmatrix} 0_{k(y)\times k(f)} & 0_{k(y)\times f} & 0_{k(y)\times f} & \cdots & 0_{k(y)\times f} \\ 0_{y\times k(f)} & K_f & CF_2 & \cdots & 0_{y\times f} \\ 0_{y\times k(f)} & 0_{y\times f} & K_f & \cdots & 0_{y\times f} \\ & & \ddots & & \\ 0_{y\times k(f)} & 0_{y\times f} & 0_{y\times f} & \cdots & CF_2 \end{bmatrix} \\
\bar{\alpha}_{s-k} &= \bar{C}_{s-k}\alpha_{s-k} \\
&+ \begin{bmatrix} 0_{k(y)\times k(f)} & 0_{k(y)\times f} & \cdots & 0_{k(y)\times f} & 0_{k(y)\times f} \\ 0_{y\times k(f)} & CF_1 & \cdots & 0_{y\times f} & 0_{y\times f} \\ 0_{y\times k(f)} & 0_{y\times f} & \cdots & 0_{y\times f} & 0_{y\times f} \\ & & \ddots & & \\ 0_{y\times k(f)} & 0_{y\times f} & \cdots & CF_1 & 0_{y\times f} \end{bmatrix}.
\end{aligned}
$$

Based on the defined matrices ϕ, V, α, β, $\bar{C}$, ζ and $\bar{\alpha}$, the mean of residual within the evaluation window can be expressed as a function of the estimation error and faults. For the convenience of expression, the residual in the evaluation window, the estimation error at the boundaries of the evaluation window and the faults in the evaluation window are described by stacking the variables at all point into column vectors, i.e.,

$$\bar{\mathbf{r}}(i, i-s+1; j, j-t) = \begin{bmatrix} r(i,j) \\ r(i-1,j) \\ \vdots \\ r(i-s+1,j) \\ r(i,j-1) \\ r(i-1,j-1) \\ \vdots \\ r(i-s+1,j-1) \\ \vdots \\ r(i,j-t) \\ \vdots \\ r(i-s+1,j-t) \end{bmatrix},$$

$$\mathbf{f}(i, i-s; j, j-t) = \begin{bmatrix} f(i,j) \\ f(i-1,j) \\ \vdots \\ f(i-s,j) \\ f(i,j-1) \\ f(i-1,j-1) \\ \vdots \\ f(i-s,j-1) \\ \vdots \\ f(i,j-t) \\ \vdots \\ f(i-s,j-t) \end{bmatrix},$$

and

$$\mathbf{e}(i, i-s+1; j-t+1) = \begin{bmatrix} e(i, j-t+1) \\ e(i-1, j-t+1) \\ \vdots \\ e(i-s+1, j-t+1) \end{bmatrix},$$

$$\mathbf{e}(i-s; j, j-t) = \begin{bmatrix} e(i-s, j) \\ e(i-s, j-1) \\ \vdots \\ e(i-s, j-t) \end{bmatrix}.$$

The relation of the residual with the estimation error and the faults can then be obtained:

$$\begin{aligned} \bar{\mathbf{r}}(i, i-s+1; j, j-t) = \ & \Psi_1 \mathbf{e}(i, i-s+1; j-t+1) \\ & + \Psi_2 \mathbf{e}(i-s; j, j-t) \\ & + \Psi_3 \mathbf{f}(i, i-s; j, j-t), \end{aligned} \tag{4.25}$$

where

$$\Psi_1 = \begin{bmatrix} \psi_{11} \\ \psi_{12} \\ \vdots \\ \psi_{1t} \end{bmatrix}, \qquad \Psi_2 = \begin{bmatrix} \psi_{21} \\ \psi_{22} \\ \vdots \\ \psi_{2t} \end{bmatrix}, \qquad \Psi_3 = \begin{bmatrix} \psi_{31} \\ \psi_{32} \\ \vdots \\ \psi_{3t} \end{bmatrix},$$

with ψ_{1l}, ψ_{2l} and ψ_{3l}, $l = 1, \cdots, t-1$, expressed as:

$$\psi_{1l} = \begin{bmatrix} \bar{C}_s V_s^{t-l} \\ \bar{C}_{s-1} V_{s-1}^{t-l} \\ \vdots \\ \bar{C}_1 V_1^{t-l} \end{bmatrix},$$

$$\psi_{2l} = \begin{bmatrix} 0_{y\times(l-1)(x)} & CA_2^s & \bar{C}_s \omega_{s,t-l} \\ 0_{y\times(l-1)(x)} & CA_2^{s-1} & \bar{C}_{s-1}\omega_{s-1,t-l} \\ & \ddots & \\ 0_{y\times(l-1)(x)} & CA_2 & \bar{C}_1 \omega_{1,t-l} \end{bmatrix},$$

$$\psi_{3l} = \begin{bmatrix} 0_{sy\times(l-1)(s+1)(f)} & \zeta_s & \bar{\alpha}_s & \sigma_{s,t-l} \\ 0_{sy\times(l-1)(s+1)(f)} & \zeta_{s-1} & \bar{\alpha}_{s-1} & \sigma_{s-1,t-l} \\ & & \ddots & \\ 0_{sy\times(l-1)(s+1)(f)} & \zeta_1 & \bar{\alpha}_1 & \sigma_{1,t-l} \end{bmatrix},$$

and ψ_{1l}, ψ_{2l}, ψ_{3l} for $l = t$ being

$$\psi_{1t} = \begin{bmatrix} \bar{C}_s \\ \bar{C}_{s-1} \\ \vdots \\ \bar{C}_1 \end{bmatrix},$$

$$\psi_{2t} = \begin{bmatrix} 0_{y\times(t-1)(x)} & CA_2^s \\ 0_{y\times(t-1)(x)} & CA_2^{s-1} \\ & \ddots \\ 0_{y\times(t-1)(x)} & CA_2 \end{bmatrix},$$

$$\psi_{3t} = \begin{bmatrix} 0_{sy\times(t-1)(s+1)(f)} & \zeta_s & \bar{\alpha}_s \\ 0_{sy\times(t-1)(s+1)(f)} & \zeta_{s-1} & \bar{\alpha}_{s-1} \\ & \ddots & \\ 0_{sy\times(t-1)(s+1)(f)} & \zeta_1 & \bar{\alpha}_1 \end{bmatrix}.$$

Equation (4.25) shows that the mean of the residual signal is directly related with the estimation error at the boundaries of the evaluation window as well as the fault value within the evaluation window. Since the mean of the estimation error at the boundaries of the evaluation window is also caused by faults occurring outside the evaluation window, the mean of residual signal within the evaluation window is solely determined by faults. It is, therefore, reasonable to detect faults by evaluating the residual signal over the evaluation window.

4.2.3 Residual Evaluation

Fault detection can be carried out by evaluating those faults whose energy level is higher than a tolerant limit L_f, i.e.,

$$\begin{aligned} \|f(i,j)\|_{s,t} &= \sqrt{\frac{\mathbf{f}(i,i-s;j,j-t)^T\mathbf{f}(i,i-s;j,j-t)}{(s+1)(t+1)}} \\ &\begin{cases} \leq L_f, & \text{fault free,} \\ > L_f, & \text{fault.} \end{cases} \end{aligned} \tag{4.26}$$

Assume that the estimation error $\mathbf{e}(i,i-s+1;j-t+1)$ and $\mathbf{e}(i-s;j,j-t)$ at the boundaries of the evaluation window in (4.25) are small enough such that

$$\bar{\mathbf{r}}(i,i-s+1;j,j-t) \approx \Psi_3\mathbf{f}(i,i-s;j,j-t). \tag{4.27}$$

The generalized likelihood ratio (GLR) is calculated for the model (4.27) and fault detection is performed using the evaluation function (Ding, 2008):

$$J_f = \mathbf{r}(i,i-s+1;j,j-t)^T(\Psi_3\Psi_3^T)^{-1}\mathbf{r}(i,i-s+1;j,j-t). \tag{4.28}$$

The decision on fault occurrence can thus be made based on:

$$J_f \leq \bar{L}_f^2, \text{ fault free,}$$
$$J_f > \bar{L}_f^2, \text{ fault,}$$

where $\bar{L}_f = (s+1)(t+1)L_f$.

If the assumption in (4.27) is removed, i.e., $\mathbf{e}(i, i-s+1; j-t+1)$ and $\mathbf{e}(i-s; j, j-t)$ are not negligible, $\mathbf{e}(i, i-s+1; j-t+1)$ and $\mathbf{e}(i-s; j, j-t)$ can be the considered as the components of faults as they are driven by faults. Define:

$$\begin{aligned} \bar{\mathbf{f}}(i, i-s; j, j-t) &= \begin{bmatrix} \mathbf{e}(i, i-s+1; j-t+1) \\ \mathbf{e}(i-s; j, j-t) \\ \mathbf{f}(i, i-s; j, j-t) \end{bmatrix}, \\ \Psi &= \begin{bmatrix} \Psi_1 & & \\ & \Psi_2 & \\ & & \Psi_3 \end{bmatrix}. \end{aligned} \tag{4.29}$$

From (4.25), the residual signal over the evaluation window is then expressed as:

$$\bar{\mathbf{r}}(i, i-s+1; j, j-t) = \Psi \bar{\mathbf{f}}(i, i-s; j, j-t). \tag{4.30}$$

Fault detection can be performed as:

$$\begin{aligned} &\|\bar{r}(i,j)\|_{s,t} \\ &\quad = \sqrt{\frac{\bar{\mathbf{r}}(i, i-s+1; j, j-t)^T \bar{\mathbf{r}}(i, i-s+1; j, j-t)}{s(t+1)}} \\ &\quad \begin{cases} \leq L_{\bar{r}}, & \text{fault free,} \\ > L_{\bar{r}}, & \text{fault,} \end{cases} \end{aligned} \tag{4.31}$$

where $L_{\bar{r}}$ is a constant determined by:

$$L_{\bar{r}} = \|C[I - (\bar{A}_1 z_1 + \bar{A}_2 z_2)]^{-1}(\bar{F}_1 z_1 + \bar{F}_2 z_2)\|_\infty L_f.$$

Defining the evaluation function as:

$$J_r = \mathbf{r}(i, i-s+1; j, j-t)^T \mathbf{r}(i, i-s+1; j, j-t), \tag{4.32}$$

the decision rule on fault detection becomes:

$$J_r \le \bar{L}_{\tilde{r}}^2, \text{ fault free,}$$
$$J_r > \bar{L}_{\tilde{r}}^2, \text{ fault,}$$

where $\bar{L}_{\tilde{r}}^2 = s(t+1)L_{\tilde{r}}^2$ and J_r being the test statistic. In online fault detection, whenever there is new measurement available, the residual value is obtained as described in Section 2.1, and J_r is calculated based on the residual from (4.32). Fault can then be detected by comparing J_r with a threshold value.

To set the threshold J_{th}, the following statistic value is introduced:

$$\tilde{\mathbf{r}}(i, i-s+1; j, j-t) = \Sigma_r^{-1/2}\mathbf{r}(i, i-s+1; j, j-t),$$

where Σ_r is the variance of residual within the evaluation window. $\tilde{\mathbf{r}}(i, i-s+1; j, j-t)$ is non-centrally χ^2 distributed with non-central parameter $\bar{\delta}_r^2$ and $\bar{\delta}_r^2 = \gamma L_{\tilde{r}}^2$ with γ = $\lambda_{min}(\Sigma_r)$ denoting the minimal eigenvalue of Σ_r. To obtain the threshold J_{th}, it is necessary to solve the threshold $\tilde{J}_{th}$ corresponding to $\tilde{\mathbf{r}}$ is solved using

$$\text{prob}\left(\chi^2(\dim(\tilde{\mathbf{r}}(i, i-s+1; j, j-t), \bar{\delta}_r^2) > \tilde{J}_{th}\right) = \alpha,$$

where $\dim(\tilde{\mathbf{r}}(i, i-s+1; j, j-t)$ is the degrees of freedom in $\tilde{\mathbf{r}}$ and α is the allowable fault alarm rate. The threshold J_{th} for J_r in (4.32) can thus be obtained:

$$J_{th} = \gamma \tilde{J}_{th}.$$

In on-line fault detection of 2-D systems, when there is new measurement available, fault detection can be performed by generating the residual signal, calculating the testing statistic J_r, and comparing J_r with the obtained threshold J_{th}. The proposed fault detection method does not require demanding on-line computation and is applicable to a wide range of 2-D systems described by the F-M model.

4.3 Simulations

The proposed Kalman filter based fault detection method is evaluated using two examples. The evolvement of residual signals in these two examples are examined to determine if it would reflect the occurrence of faults.

Example 1. In this example, the following system with $x(i,j) \in \mathbb{R}^2$ is considered:

$$\begin{aligned}
x(i+1,j+1) &= \begin{bmatrix} 0.1 & 0.2 \\ 0 & 0.2 \end{bmatrix} x(i+1,j) + \begin{bmatrix} 0.1 & 0 \\ 0.7 & 0.2 \end{bmatrix} x(i,j+1) \\
&\quad + \begin{bmatrix} 0.1 \\ 0.2 \end{bmatrix} u(i+1,j) + \begin{bmatrix} 0 \\ 0.1 \end{bmatrix} u(i,j+1) \\
&\quad + \begin{bmatrix} 0.2 \\ 0 \end{bmatrix} \eta(i+1,j) + \begin{bmatrix} 0.1 \\ 0.1 \end{bmatrix} \eta(i,j+1) \\
&\quad + \begin{bmatrix} 0.1 \\ 0.2 \end{bmatrix} f(i+1,j) + \begin{bmatrix} 0 \\ 0.1 \end{bmatrix} f(i,j+1) \\
y(i,j) &= \begin{bmatrix} 0.1 & 0.1 \end{bmatrix} x(i,j) + 0.03 f(i,j) + 0.05 \mu(i,j)
\end{aligned} \tag{4.33}$$

In the simulation, the spacial index i ranges from 1 to 50. The system noise η and ν are also random variables ranging from 0 to 1. Figure 4.3 displays the residual signal evolvement when the input $u(i,j) = 0$ and a unit step fault signal occurs at time $j = 100$. It is observed that the residual signal is of zero mean with a small variance when no fault occurs before j=100, indicating that the state estimate from the Kalman filter is consistent with the true state when there is no fault. When a step fault occurs at time $j = 100$, the residual becomes a signal of non-zero mean with increased variance, indicating that the generated residual signal is capable of detecting fault occurrence. The increased variance during the fault occurrence implies a higher evaluation function value in (4.28).

The input signal u is then set to be a random variable varying between 0 and 1 and the generated residual is examined when a unit step fault signal occurs at time $j = 100$. By setting

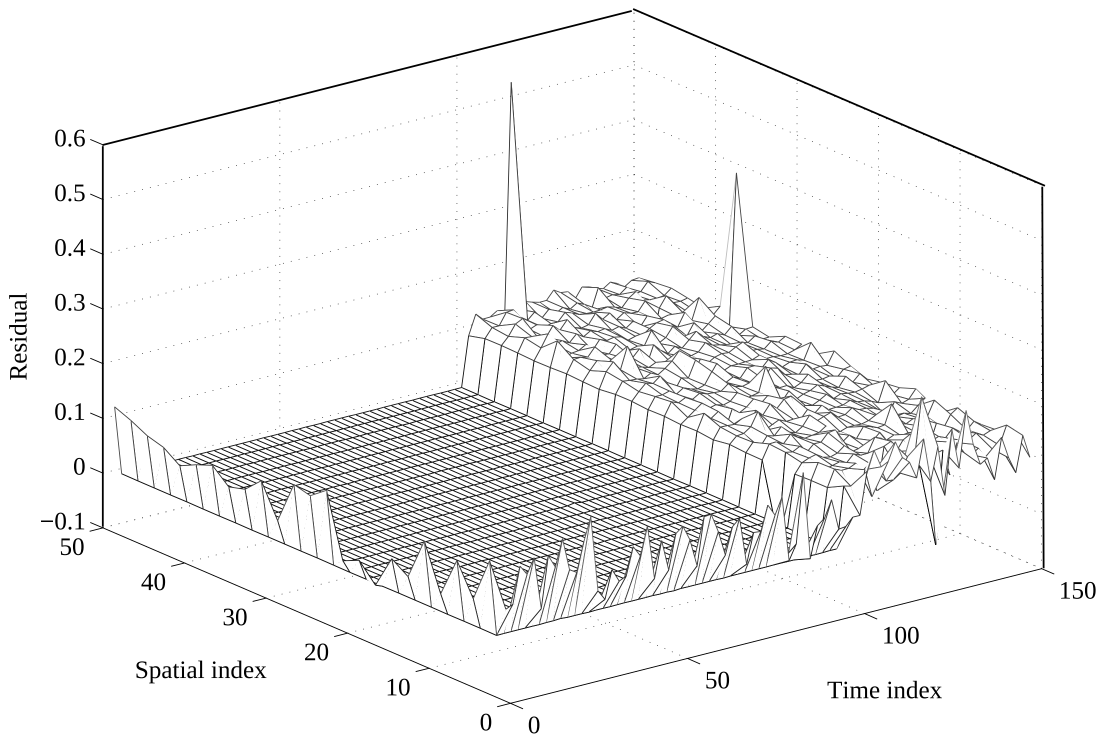


Figure 4.3: 3-D plot of residual evolvement when a unit step fault occurs at j=100 and u=0 in Example 1.

the tolerable fault energy $L_f = 0.05$, the evaluations window as 6×3 and the fault alarm rate $\alpha = 0.05$, the threshold value is calculated to be $J_{th} = 0.017$. The testing statistic J_r can then be evaluated in comparison with the threshold value. Figure 4.4 and shows the responses of the residual r and J_r at $i = 12,\ 22$ and 32 to the step change in f. When there is no fault, the residual r and j_r are very small, implying that the state estimate converges to the true state after a very short time using the 2-D Kalman filter. When the step fault occurs, r increases and settles at a non-zero value with an increased variance and r displays a non-zero mean value. The residual evaluation J_r remains within the control limit when there is no fault and it increase and violate the threshold value when the fault occurs. This indicates that the occurrence of a fault can trigger changes in r and J_r and that the proposed fault detection strategy is effective for this system. The 3-D plot of residual in Figure 4.5 displays that the residual responds to the fault across the spatial direction at $j = 100$ and residual changes from zero-mean to non-zero mean when the step fault occurs. The 3-D plot of residual evaluation J_r in Figure 4.6 indicates that the residual evaluation J_r remains within the control limit when there is no fault and it exceeds the threshold value when the fault occurs. The Kalman filter based fault detection can be used to detect faults no matter whether there is a process input signal or not.

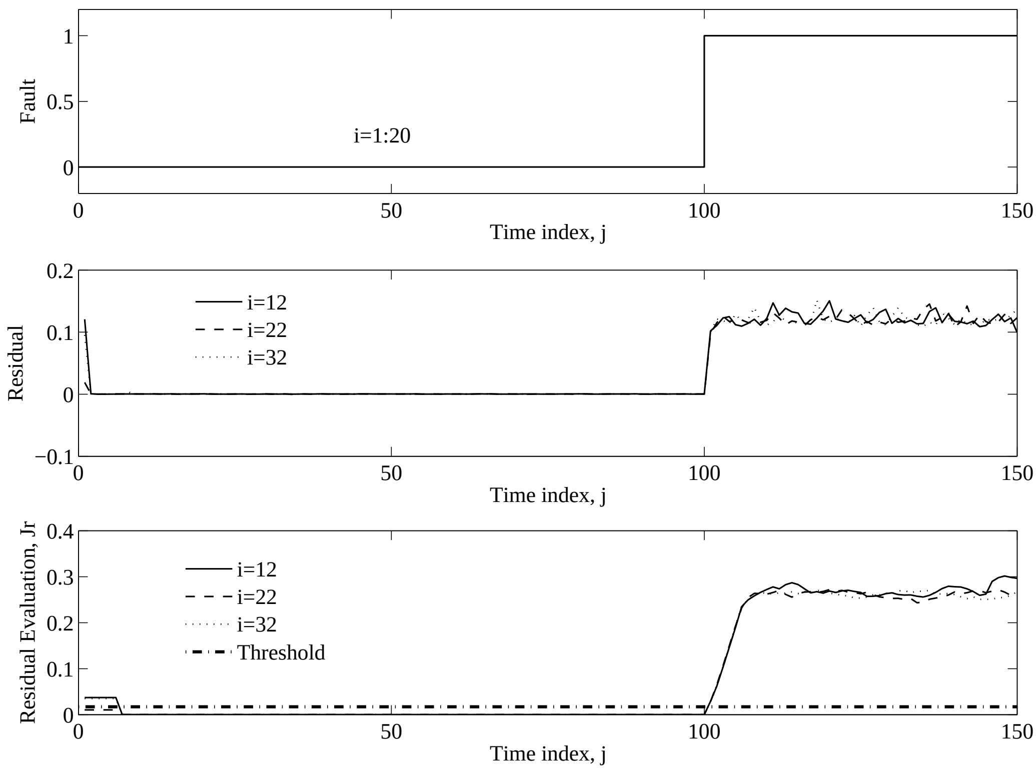


Figure 4.4: Residual and residual evaluation at different spatial points when a unit step fault occurs at j=100 and u is a random signal in Example 1 .

To investigate the effect of different fault signal on the performance of the fault detection method, a sinusoidal fault is introduced to the system at $j = 100$ as in Figure 4.7. Note that the sinusoidal signal used has a phase shift at each spatial point from its neighboring point. The residual response to the sinusoidal fault is shown 4.8. It is observed that the residual response to a fault signal of zero mean value such as a sinusoidal fault is different from that of a step fault in that the mean value of the residual signal remains to be zero when the fault occurs. The variance of the residual increases significantly when the fault occurs. As there is no observable

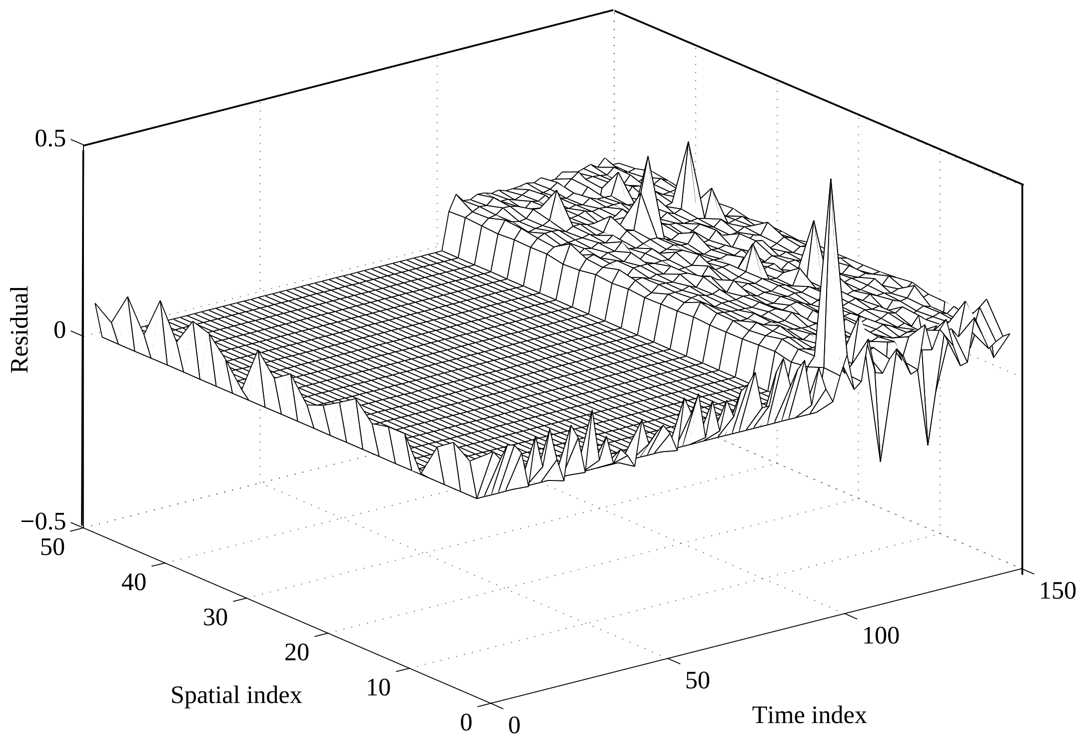


Figure 4.5: 3-D plot of residual evolvement when a unit step fault occurs at j=100 and u is a random signal in Example 1.

change in the mean of the residual, it is more convenient to examine the residual evaluation J_r for fault detection in this case. Figure 4.9 displays the 3-D plot of residual evaluation J_r for the sinusoidal fault. It indicates that the residual evaluation J_r exceeds the threshold value at all spatial points when the fault occurs. The responses of residual r and residual evaluation J_r at some spatial points are shown in Figure 4.10. The residual r responds to the fault of zero mean quickly by displaying a larger variance and the residual evaluation can be conveniently used to detect the occurrences of faults.

Example 2. In this example, a more complicated system than Example 1 with $x(i,j) \in \mathbb{R}^3$

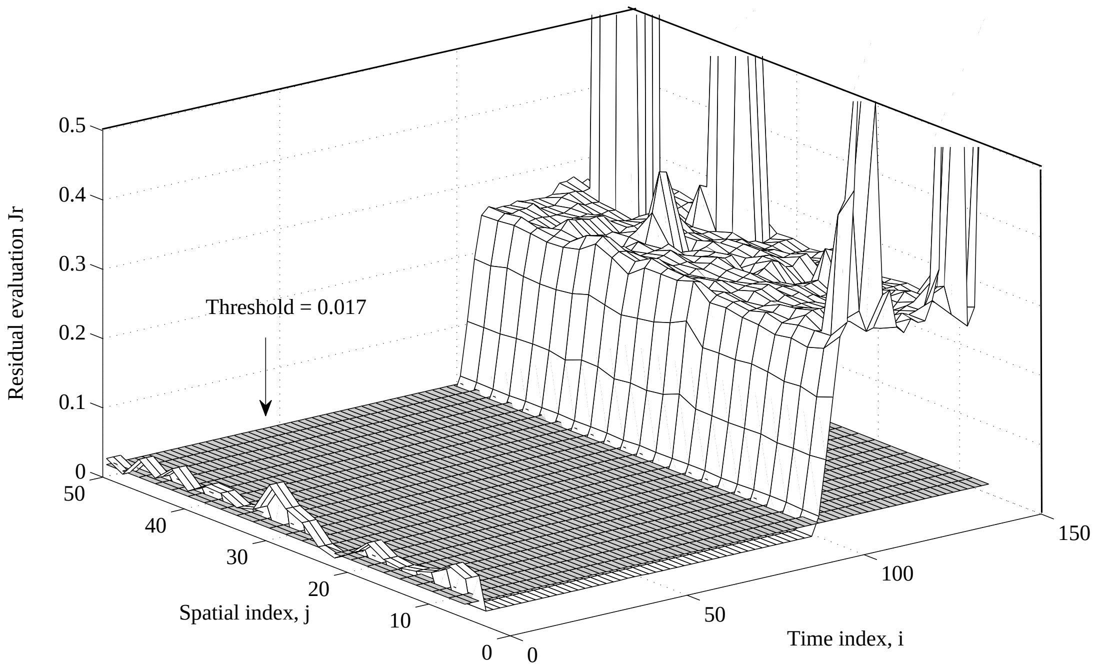


Figure 4.6: 3-D plot of residual evaluation when a unit step fault occurs at j=100 and u is a random signal in Example 1.

is considered:

$$
\begin{aligned}
x(i+1,j+1) = &\begin{bmatrix} 0.1 & 0.2 & 0 \\ 0 & 0.2 & 0.1 \\ 0 & 0.1 & 0 \end{bmatrix} x(i+1,j) + \begin{bmatrix} 0.1 & 0 & 0.3 \\ 0.2 & 0.2 & 0 \\ 0.2 & 0.4 & 0 \end{bmatrix} x(i,j+1) \\
&+ \begin{bmatrix} 0.5 \\ 0 \\ 0.1 \end{bmatrix} u(i+1,j) + \begin{bmatrix} 0 \\ 0.3 \\ 0.2 \end{bmatrix} u(i,j+1) + \begin{bmatrix} 0.1 \\ 0 \\ 0 \end{bmatrix} f(i+1,j) \\
&+ \begin{bmatrix} 0 \\ 0.3 \\ 0 \end{bmatrix} f(i,j+1) + \begin{bmatrix} 0.2 \\ 0.1 \\ 0 \end{bmatrix} \eta(i+1,j) + \begin{bmatrix} 0 \\ 0.1 \\ 0.1 \end{bmatrix} \eta(i,j+1) \\
y(i,j) = &\begin{bmatrix} 0 & 0.3 & 0.2 \end{bmatrix} x(i,j) + 0.03 f(i,j) + 0.05\mu(i,j)
\end{aligned}
\tag{4.34}
$$

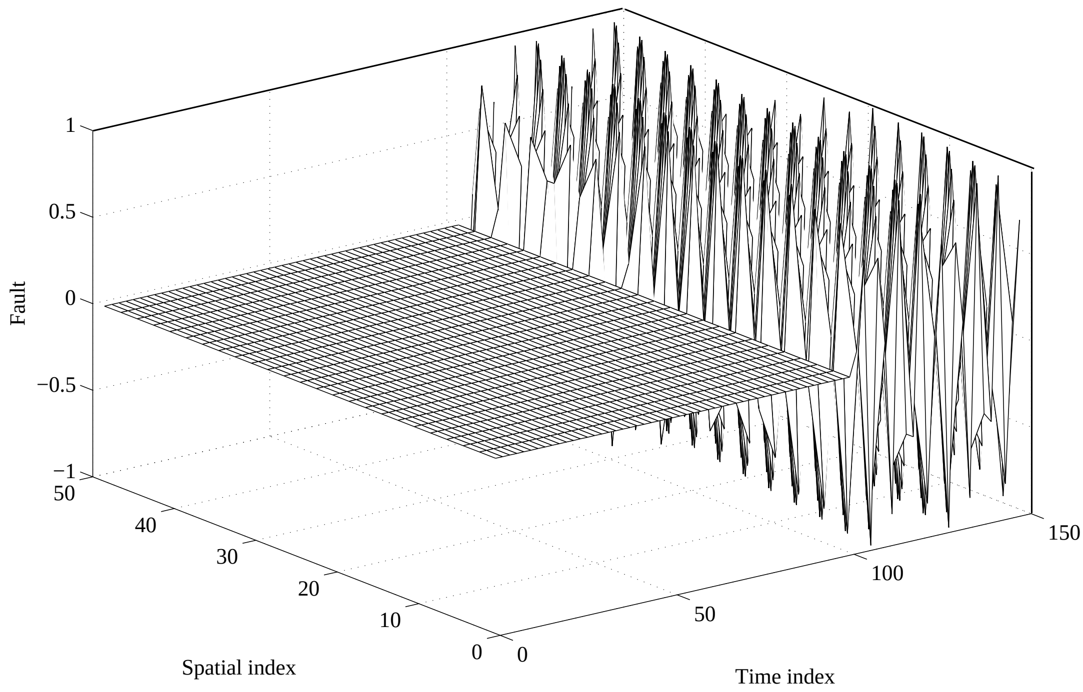


Figure 4.7: 3-D plot of a sinusoidal fault signal.

Simulation was carried out to evaluate the effectiveness of the fault detection strategy for this system. In the simulation, the spacial index i ranges from 1 to 20, the input signal u is set to be a random variable varying from 0 to 1 and the noise signals η and μ in (4.34) are random variables varying from 0 to 1. In residual evaluation, the tolerable fault limit is set to be $L_f = 0.05$ and with the evaluation window being 6×3, the threshold value is calculated to be 0.105. The residual r and residual evaluation J_r are examined to determine whether a fault has occurred. Figure 4.11 displays the 3-D plot of the residual response to a unit step fault occurring at $j = 100$. It can be seen that the residual signal r is zero mean before $j = 100$ and becomes non-zero mean after a fault occurs, although the non-zero mean is not as obvious as that in Figure 4.4 and Figure 4.5 of Example 1 due to a much higher variance. The much increased variance and some large spikes when a fault occurs is due to the complexity of the system and effect of couple noises. The 3-D plot of residual evaluation J_r is shown in Figure 4.12. The residual evaluation is within the threshold limit when there is no fault and it exceeds

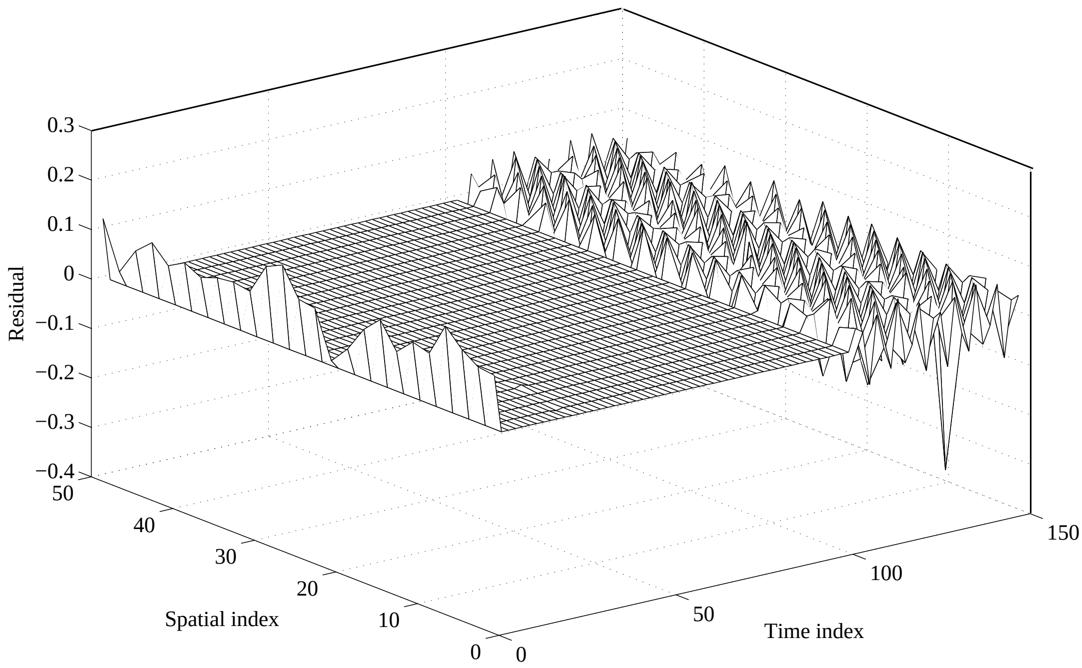


Figure 4.8: 3-D plot of residual evolvement when a sinusoidal fault occurs at j=100 and u is a random signal in Example 1.

the threshold value when the fault occurs. The large variances and large spikes in the residual r also lead to the large variances in the residual evaluation J_r. Variations of the residual r and the residual evaluation J_r at several spatial points can be clearly examined in Figure 4.13 and the fault can be detected by examining the residual and the residual evaluation signal.

The proposed fault detection method is also evaluation for a sinusoidal fault occurring at $j = 100$. As in 4.14, for a fault of zero mean, the mean value of the residual signal does not have any noticeable change when the fault occurs, and the residual r responds to the fault with an increased variance. In comparison with that in Example 1, the residual displays a larger variances and more sharp spikes due to the increased complexity of the system. From Figure 4.15, it can be seen that the residual evaluation J_r is within the control limit and it violates the threshold when the fault occurs. Figure 4.16 shows the response of the residual and the evaluation function J_r at several spatial points for the sinusoidal fault. It is clear that the

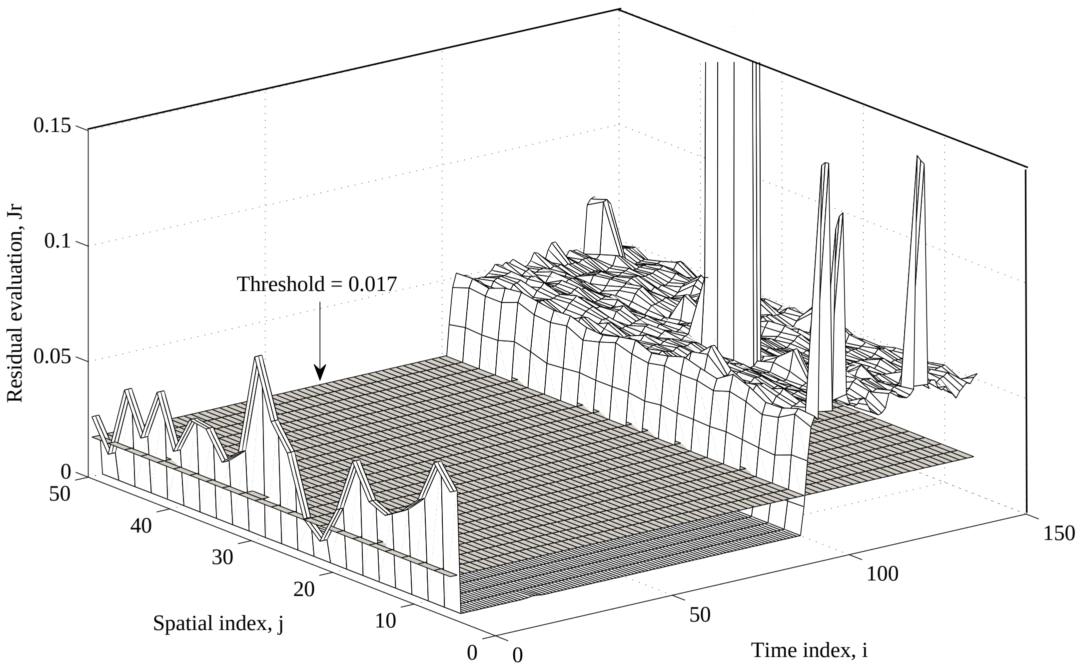


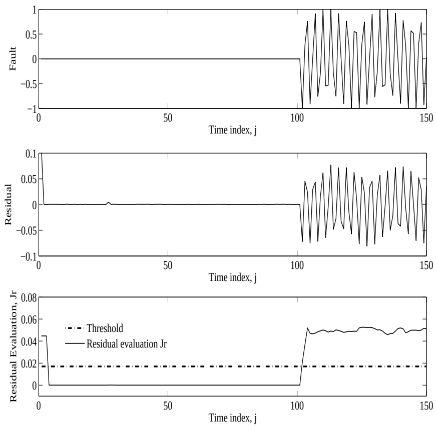
Figure 4.9: 3-D plot of residual evaluation when a sinusoidal fault occurs at j=100 and u is a random signal in Example 1.

evaluation function J_r can be used to determine whether a fault occurs.

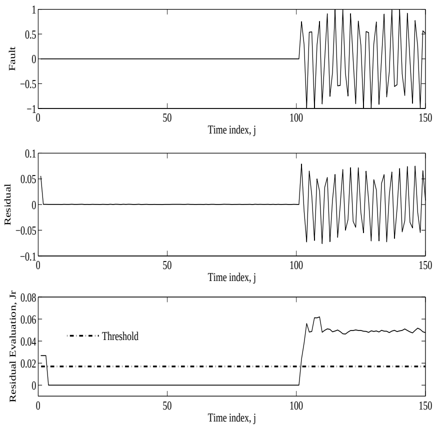
4.4 Summary

Many systems contain stochastic noises that cannot be decoupled in a residual generator. Fault detection of such systems has been developed in one-dimensional domain. Fault detection of 2-D systems is a challenging subject and only limited results are available. In this chapter, a fault detection strategy is developed for 2-D systems containing stochastic noises. Based on the 2-D F-M models, a recursive Kalman filter is used to derive a state estimate minimizing the variances of state estimate error. The obtained state estimate is then compared with the measured information, leading to generation of a residual signal. Since the residual signal reflects the fault information, it can be evaluated to determine whether a fault occurs. Due to the stochastic characteristics of the system, it is reasonable to evaluate the residual over an evaluation window. In 2-D system, the evaluation window is defined by a moving rectangular plane with the limits of two independent variables corresponding to two directions. A model relating the residual to faults over the evaluation window is derived, from which the generalized likelihood test technique is then applied to determine if a fault has occurred.

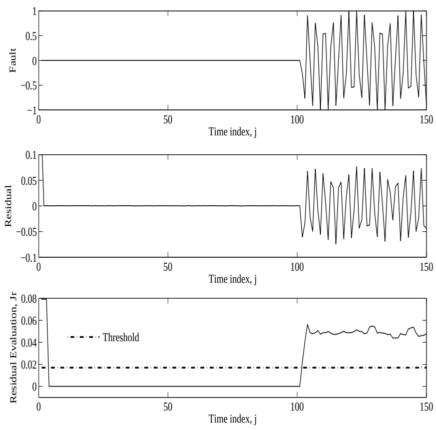
The proposed fault detection method is evaluated using two examples via simulations. Simulation results show that the residual is of zero mean with a variance when there exists no fault. When a step fault occurs, the residual responds quickly and evolves to a signal of non-zero mean with increased variances. For a fault of zero mean, such as a sinusoidal fault, the residual responds to the fault with increased variances although there is no noticeable change in its mean value. For both the step and sinusoidal faults, the evaluation function remains within the control limit when there is no fault and violate the threshold what the fault occurs. It can be used to clearly determine the occurrences of faults. The proposed fault detection method is able to detect faults effectively and efficiently. This method does not require to satisfy strict conditions in order for a residual generator to exist and can, therefore, be applicable to a general 2-D systems described by F-M models.



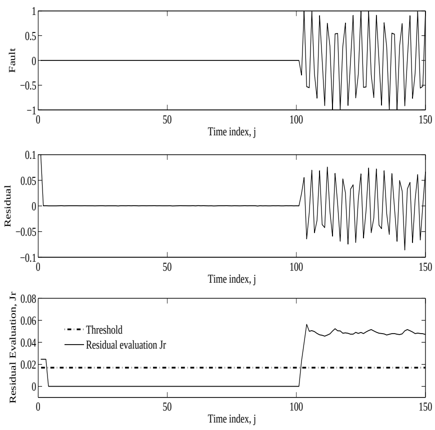
(a) i=10



(b) i=20



(c) i=30



(d) i=40

Figure 4.10: Residual and residual evaluation at different spatial points when a sinusoidal fault occurs at j=100 and u is a random signal in Example 1.

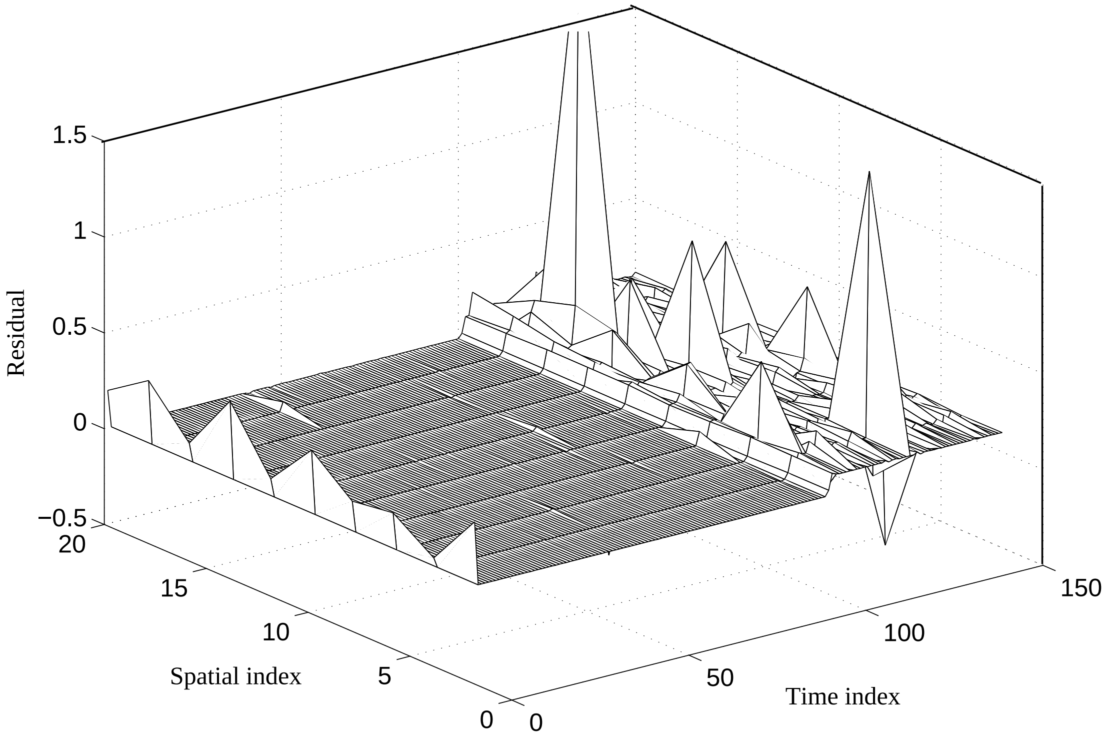


Figure 4.11: 3-D plot of residual evolvement when a unit step fault occurs at j=100 and u is a random signal in Example 2.

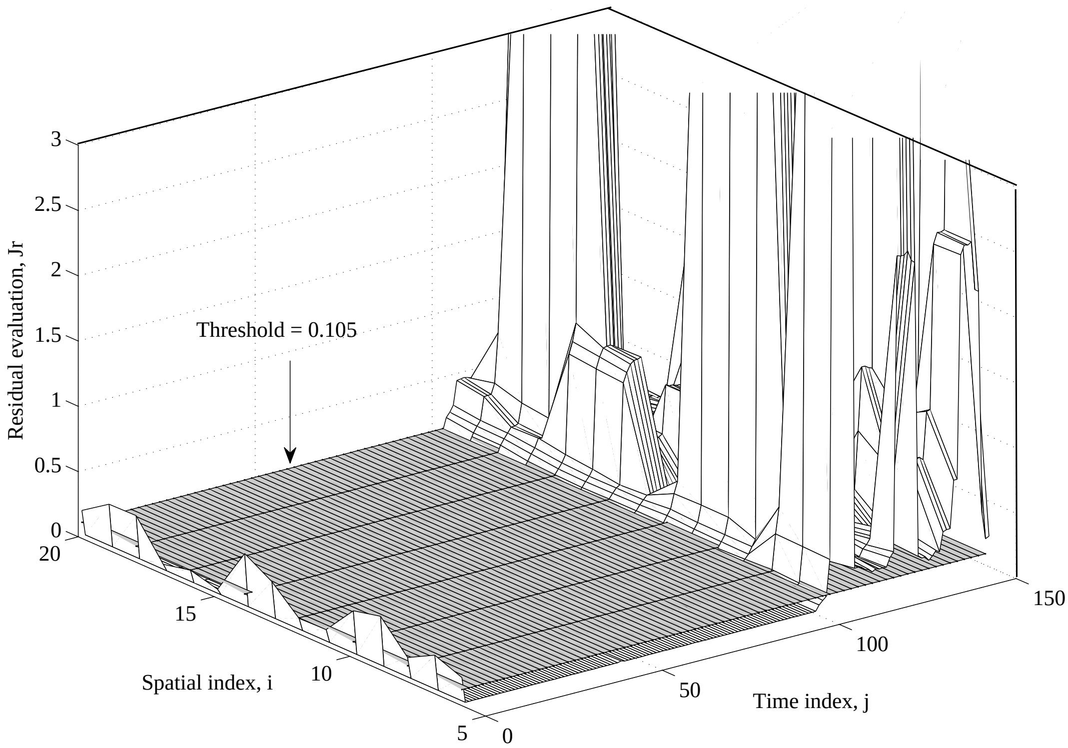


Figure 4.12: 3-D plot of residual evaluation when a unit step fault occurs at j=100 and u is a random signal in Example 2.

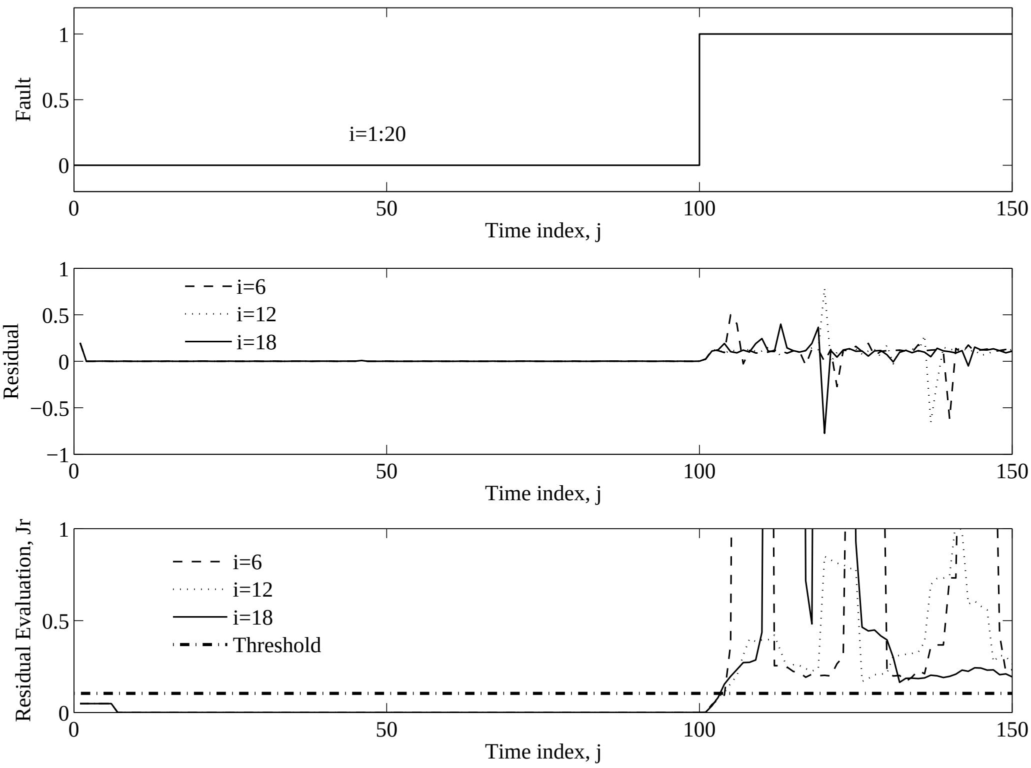


Figure 4.13: Residual and residual evaluation at different spatial points when a unit step fault occurs at j=100 and u is a random signal in Example 2.

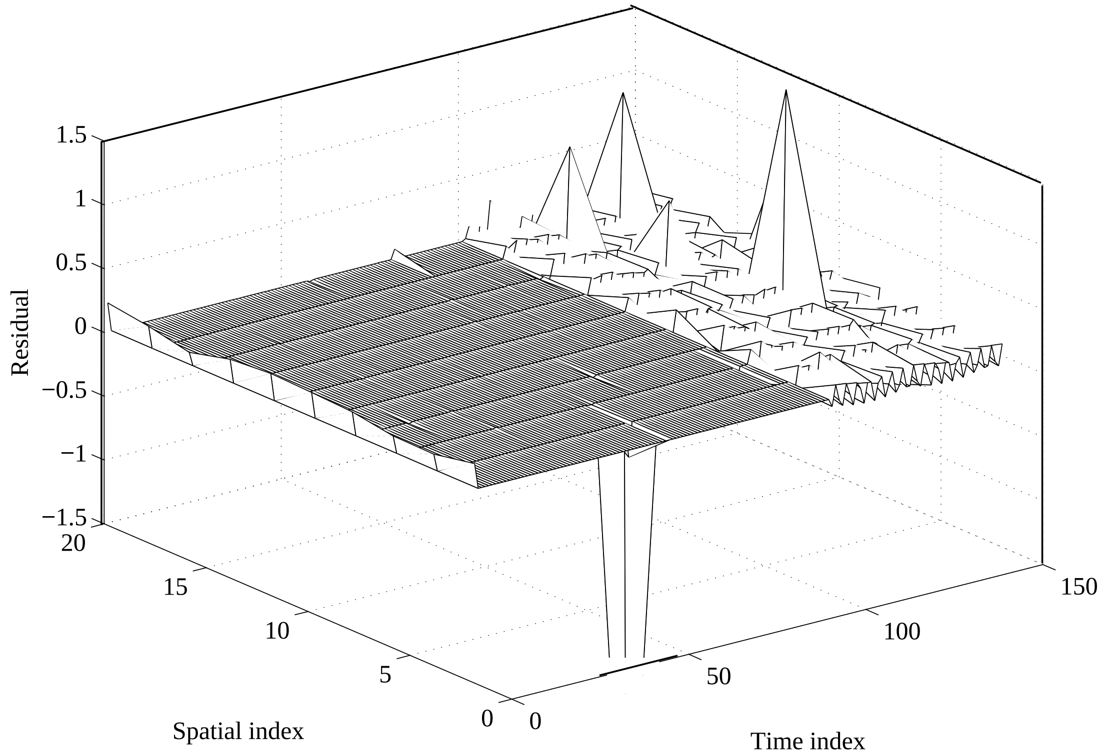


Figure 4.14: 3-D plot of residual evolvement when a sinusoidal fault occurs at j=100 and u is a random signal in Example 2.

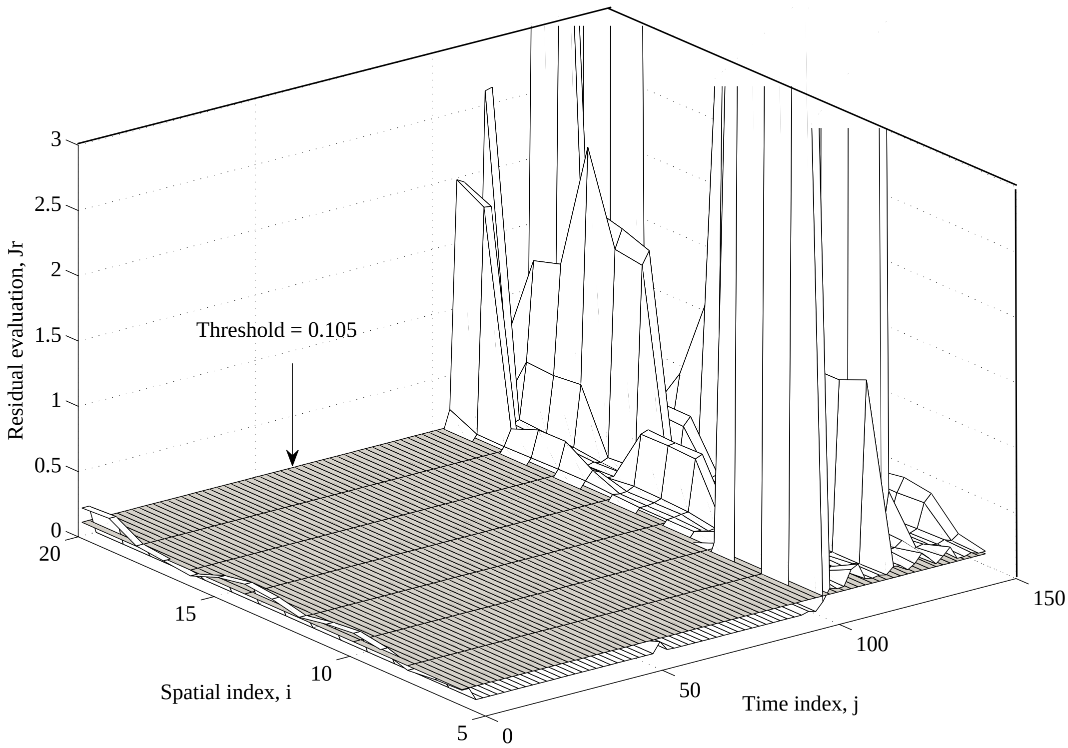
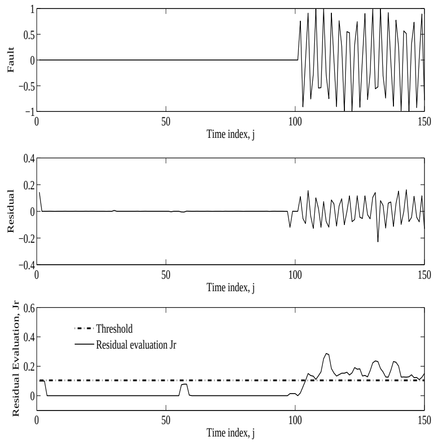
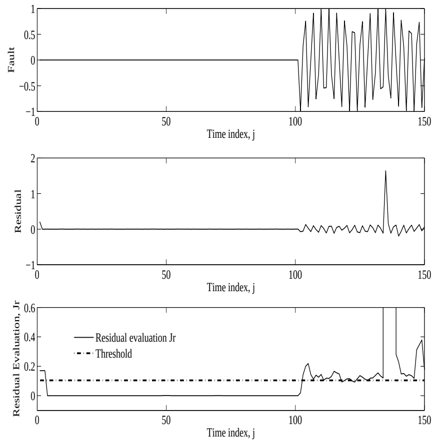


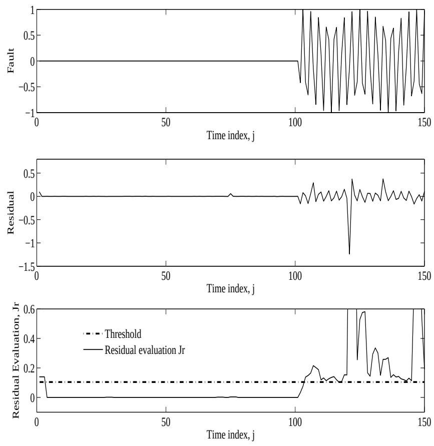
Figure 4.15: 3-D plot of residual evaluation when a sinusoidal fault occurs at j=100 and u is a random signal in Example 2.



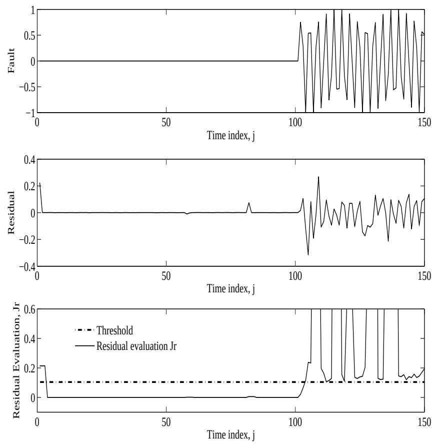
(a) i=6



(b) i=10



(c) i=15



(d) i=20

Figure 4.16: Residual and residual evaluation at different spatial points when a sinusoidal fault occurs at j=100 and u is a random signal in Example 2.

Chapter 5

Fault Detection of Two Dimensional Systems Using DPCA

5.1 Introduction

Principal component analysis (PCA) is a well-known data dimensionality reduction technique that has been used in system monitoring, diagnosis and control (Yoon and MacGregor, 2001). In the conventional PCA, it is assumed that processes operate in a "steady-state" wherein the originating signals can be considered as statistically stationary - the signal properties do not change with time. Dynamic principal component analysis (DPCA) takes into account serial correlations in the data by augmenting each observation vector with the previous d observations, and has been applied in detecting faults in dynamic processes and batch operations wherein time variations are significant (Russell et al., 2000b; Doan and Srinivasan, 2008). Existing applications of PCA and DPCA have been focused on 1-D systems.

In the last three chapters, the fault detection approaches have been proposed for 2-D systems and these approaches involve generating residual signals that reflect fault information. The observer based fault detection method for systems modeled by Roesser models consists of development of an observer and a residual generator. In the Kalman filter based fault detection method, residual signals are generated by comparing the measurement with state estimate obtained from a recursive Kalman filter. The generated residual signals can reflect the dynamic change of faults such as step fault signals, and faults can, therefore, be detected by examining the variations of residual curves. In many cases, however, the fault may be irregular and associated with other variables, and the residual signal may be corrupted with disturbances and uncertainties caused by parameter changes. It may not be obvious to determine whether a fault occurs simply by observing the resulting residual signals. To achieve a successful fault detection based on the available residual signal, further efforts are needed. A widely accepted way is to generate such a feature of the residual signal, by which we are able to distinguish the faults from the disturbances and uncertainties. Residual evaluation and threshold setting serve for this purpose. A decision on the possible occurrence of a fault is made by means of a simple comparison between the residual feature and the threshold. It is therefore necessary to investigate 2-D fault detection utilizing the concept of residual evaluation and threshold setting. The residual evaluation and threshold setting has been discussed for the Kalman filter based fault detection in Chapter 4. In this chapter, an approach of residual evaluation and threshold is proposed that applies to both Kalman filter based and observer based fault detection methods.

In many 2-D systems, the horizontal (or spatial) range is binary bounded and it is possible to investigate the timely evolvement of the residual signals at all spatial points simultaneously and detect faults by evaluating the evolvement of residual signals. In this chapter, fault detection is studied by considering the residual signals at all spatial points as multivariate signals. From the discussion from the previous chapters, a 2-D residual signal can be generated using the observer based fault detection or Kalman filter based fault detection. The 2-D residual signals are then evaluated at all spatial points simultaneously using DPCA and fault detection is performed using a simpler 1-D testing statistic. Two examples are used to examine the performance of the strategy, in one example the residual is generated using the observer based fault detection and in the other the residual generated using the Kalman filter.

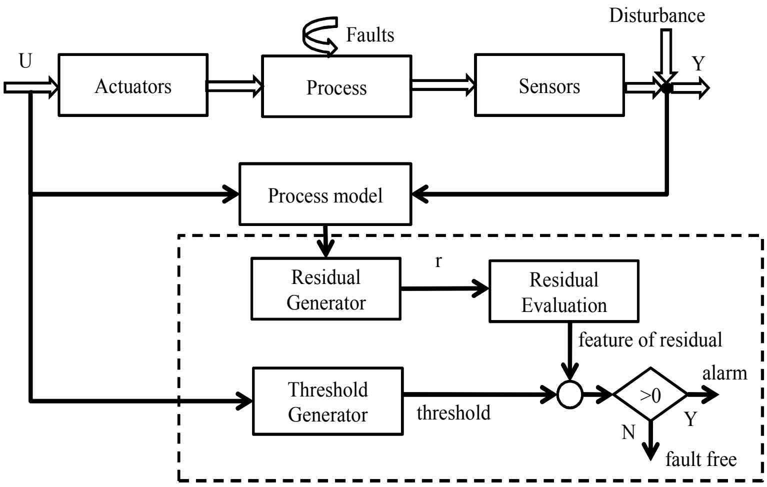


Figure 5.1: The procedure of a model-based fault detection procedure

5.2 Fault detection of 2-D systems using DPCA

In model based fault detection of 2-D systems, residual signals reflecting the dynamic change of faults are generated based on the 2-D models. The residual signals are then evaluated using DPCA to determine if a fault occurs. The schematic of the model based fault detection is shown in Figure 5.1.

5.2.1 Residual generation

2-D systems have been represented by Roesser models and F-M models. A dead-beat observer based fault detection was proposed for F-M models (Bisiaco and Valcher, 2006; Bisiacco and Valcher, 2008). In this thesis, research has been focused on observer based fault detection for Roesser models and Kalman filter based fault detection on F-M models. The observer based and Kalman filter based fault detection approaches are summarized in Figure 5.2. Both approaches involve obtaining state estimate $\hat{x}$ from process input and measured output. In the observer based fault detection, the residual generator is obtained by calculating the transfer functions and then applying the realization technique to obtain the residual $r(i,j)$. In the Kalman filter based fault detection, the residual signal is generated from comparing the state estimate with the measurement. The residual signals $r(i,j)$ from these approaches deviate from normal operations when faults occur and can therefore be used to detect faults. Although it is feasible to detect faults by simply observing the evolvement of the residual signals as discussed in the previous chapters, it is more desirable to quantitatively evaluate the residual signals and compare them with the threshold values for fault detection in order to provide more accurate fault detection, especially when the fault signals are complicated and/or the systems contain noises.

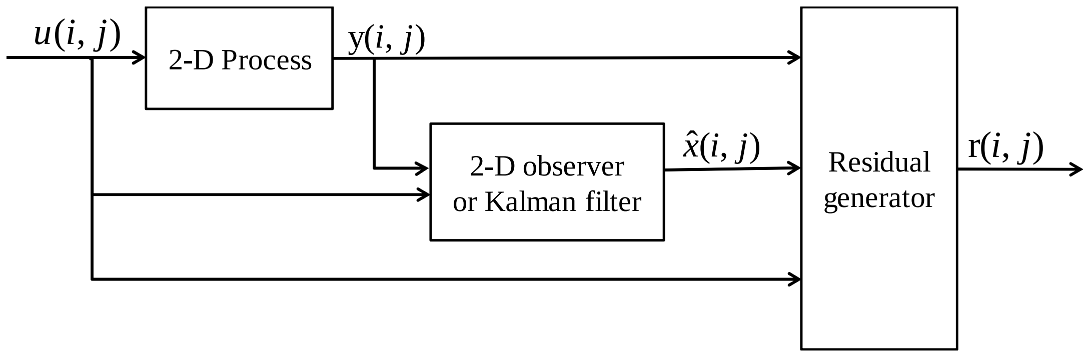


Figure 5.2: Schematic of fault detection of 2-D systems using an observer or a Kalman filter

5.2.2 Residual evaluation using DPCA

DPCA is an extension of PCA to the realm of dynamic processes with serial correlation. Some literature (Ku et al., 1995; Chen and Liu, 2002) applied it to monitor batch processes. The purpose of using DPCA is to establish a time-lagged evaluation window for off-line residual database correction. Unlike the GLR based method presented in Chapter 4, the DPCA based fault detection obtains the threshold value with a reference database during the off-line calculation. In comparison with the conventional PCA model relying on only current observation, DPCA uses the measurement within the evaluation window ranging from past to current observations, and hence can obtain more reliable detection.

Mathematically, DPCA starts with forming a time-lagged window for each of the batches in the reference database. For residual evaluation of 2-D systems, several groups of process data at normal operations are first collected and the corresponding residual signals $r(i, j)$ are generated using the observer or Kalman filter approach discussed in the previous chapters. Assume that the dimension of i and j would be M and N, and $M \leq N$. Let

$$\mathbf{r}(j) = \begin{bmatrix} r(1,j), & r(2,j), & \cdots & r(M,j) \end{bmatrix}, \; j = 1, ...N \tag{5.1}$$

The data matrix for the residual signal $\mathbf{r}(j)$ with the previous d observations can then be written as

$$R_d = \begin{bmatrix} \mathbf{r}(d+1) & \mathbf{r}(d) & \cdots & \mathbf{r}(1) \\ \mathbf{r}(d+2) & \mathbf{r}(d+1) & \cdots & \mathbf{r}(2) \\ \vdots & \vdots & & \vdots \\ \mathbf{r}(N) & \mathbf{r}(N-1) & \cdots & \mathbf{r}(N-d) \end{bmatrix} \tag{5.2}$$

The covariance matrix of the time lagged data R_d is

$$S = \frac{1}{N-d-1} R_d^T R_d \tag{5.3}$$

If I groups of data at normal operations are collected and S^i indicates the covariance matrix corresponding to the i^{th} group, the average value of S is obtained:

$$S^{\text{avg}} = \frac{(N-d-1)\sum\limits_{i=1}^{I} S^i}{I(N-d)} \tag{5.4}$$

Eigenvalues and eigenvector of the covariance matrix S^{avg} are calculated. Let p_i indicate the eigenvector corresponding to the i^{th} largest eigenvalue λ_i and the i^{th} principal component is described by $t_i = p_i^T R_d$. Corresponding to a largest eigenvalues, the data matrix R_d can be approximately represented by a principal components. Let Λ be a diagonal matrix containing the a largest eigenvalues and $P = \begin{bmatrix} p_1, & p_2, & \cdots, & p_a \end{bmatrix}$ containing the loading vectors associated with the a largest eigenvalues. The normal operation can be characterized by the Hotellings T^2 statistic

$$T^2(j) = \mathbf{r_d}(j)^T P \Lambda^{-1} P^T \mathbf{r_d}(j) \tag{5.5}$$

where $\mathbf{r_d}(j) = \begin{bmatrix} \mathbf{r}(j), & \mathbf{r}(j-1), & \cdots, & \mathbf{r}(j-d) \end{bmatrix}$. The threshold value T_α^2 is calculated using the probability distribution:

$$T^2_\alpha = \frac{a(N-d-1)(N-d+1)}{(N-d)(N-d-a)} F_{a,N-d-a,\alpha} \tag{5.6}$$

where α is the confidence limit which is generally set as 95% and $F_{a,K-d-a,\alpha}$ is the F-distribution above the critical point α with a and $N-d-a$ degrees of freedom. By calculating the residual signal $\mathbf{r}(j)$ and examining the Hotellings T^2 statistic, fault detection can be performed

$$T^2(j) \begin{cases} \leq T^2_\alpha, & \text{fault free} \\ > T^2_\alpha, & \text{fault} \end{cases} \tag{5.7}$$

Fault detection of 2-D system using DPCA is comprised of off-line and on-line procedures and can be summarized as:

- The off-line procedure includes collecting the historical reference data and calculating the residual signals of the reference data. Several groups of residual reference data are organized into the time-lagged matrix form as in (5.2) and the average covariance matrix is calculated using equation (5.4). The numbers of principal component a and the matrices P and Λ are determined using the eigenvalue decomposition. The 95% control limit T^2_α can be found by equation (5.6) and serves as the threshold setting.

- In the on-line monitoring, process input and output data are collected and the residual values are computed. The real-time T^2 for the residual signals is obtained using equation (5.5). If the value of real-time T^2 exceed the 95% control limit T^2_α, it indicates that a fault occurs. Otherwise, no fault has occurred and the system is running normally.

The procedure of fault detection using DPCA is illustrated in Figure 5.3.

5.3 Simulation

The residual evaluation using DPCA is applied to two examples, one of which is an F-M model with the residual generated from Kalman filter based approach and the other a Roesser model

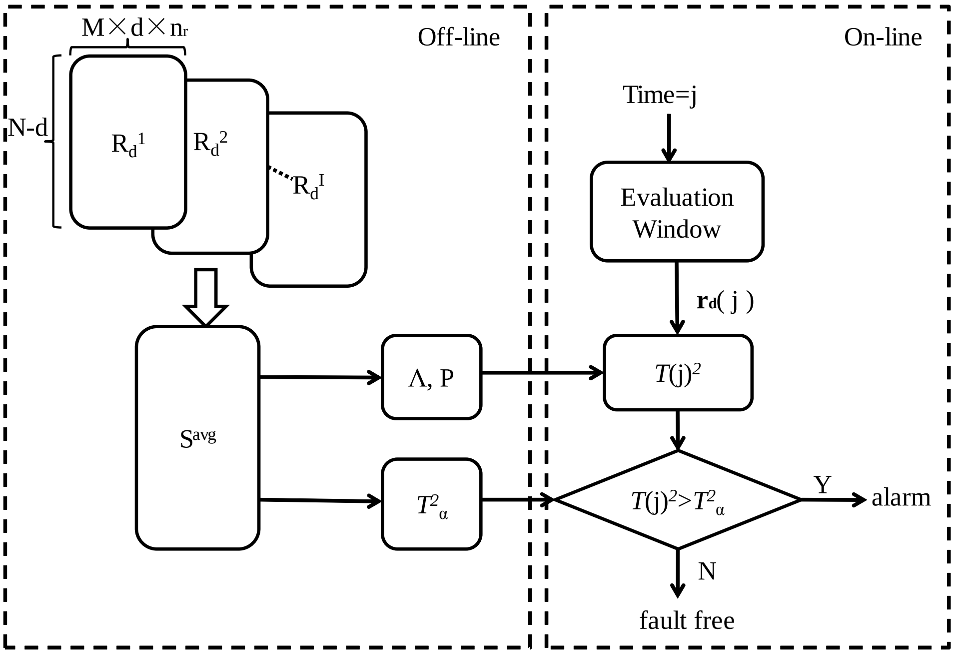


Figure 5.3: Fault isolation procedure using DPCA

with the residual from the observer based residual generator.

Example 1. In this example, the Kalman filter based fault detection using DPCA for residual evaluation is investigated. Consider the same example discussed in Chapter 4:

$$
\begin{aligned}
x(i+1,j+1) = &\begin{bmatrix} 0.1 & 0.2 \\ 0 & 0.2 \end{bmatrix} x(i+1,j) + \begin{bmatrix} 0.1 & 0 \\ 0.7 & 0.2 \end{bmatrix} x(i,j+1) \\
&+ \begin{bmatrix} 0.1 \\ 0.2 \end{bmatrix} u(i+1,j) + \begin{bmatrix} 0 \\ 0.1 \end{bmatrix} u(i,j+1) \\
&+ \begin{bmatrix} 0.2 \\ 0 \end{bmatrix} \eta(i+1,j) + \begin{bmatrix} 0.1 \\ 0.1 \end{bmatrix} \eta(i,j+1) \\
&+ \begin{bmatrix} 0.1 \\ 0.2 \end{bmatrix} f(i+1,j) + \begin{bmatrix} 0 \\ 0.1 \end{bmatrix} f(i,j+1) \\
y(i,j) = &\begin{bmatrix} 0.1 & 0.1 \end{bmatrix} x(i,j) + 0.03 f(i,j) + 0.05 \mu(i,j)
\end{aligned}
$$

A historical reference database of $I = 5$ was generated under the normal condition and the residual signals was calculated using the Kalman filter based approach described in Chapter 4. The spatial dimension $M = 50$, the evaluation window length $d = 4$ and the number of principal components $a = 8$ were used. The threshold value was calculated to be $T_\alpha^2 = 15.916$. Figure 5.4 shows the T^2 value for the residual signal when a unit step fault occurs at $j = 100$. It is observed that T^2 value is within the control limit when there is no fault. The threshold value is violated after the fault occurs and the fault is therefore detected. The response of the T^2 value to fault is almost instant and there is no noticeable delay for the step with the evaluation window $d = 4$.

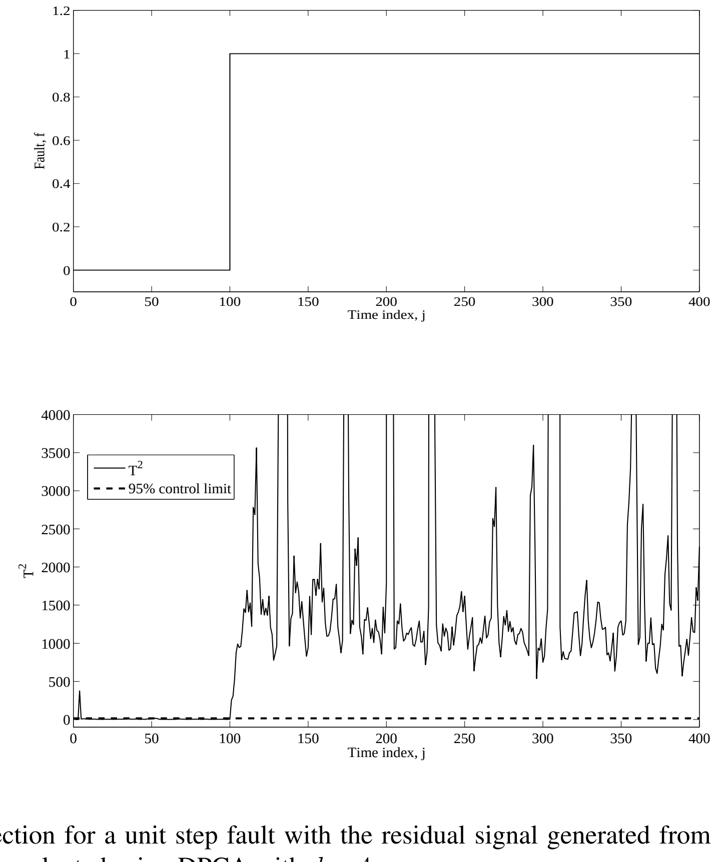


Figure 5.4: Fault detection for a unit step fault with the residual signal generated from the Kalman filter based residual generator and evaluated using DPCA with $d = 4$.

The effect of different fault signals on the proposed strategy is investigated by introducing a sinusoidal fault signal at $j = 200$. The threshold value was the same $T_\alpha^2 = 15.916$ since the same reference signals and the same evaluation window were used. Figure 5.5 shows the fault signal (at one spatial point $i = 20$) and evolvement of T^2 value. The T^2 value is smaller than the threshold value before the sinusoidal fault occurs and it exceeds the control limit after the fault occurs at $j = 200$. The indicates that the fault detection using DPCA is effective for faults with zero mean values such as sinusoidal signals.

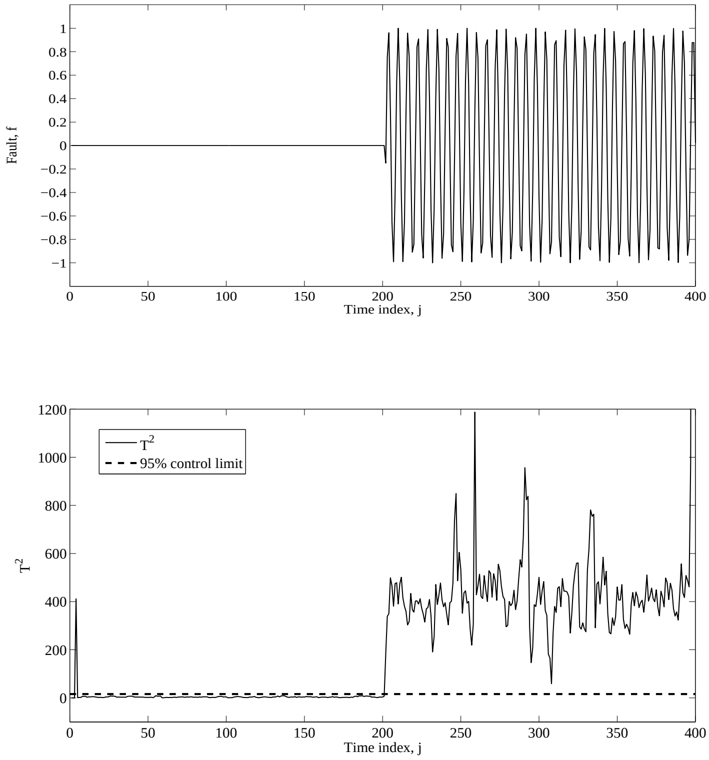


Figure 5.5: Fault detection for a sinusoidal fault with the residual signal generated from the Kalman filter based residual generator and evaluated using DPCA with $d = 4$.

To examine the effect of the length of evaluation window on the performance of the fault detection using DPCA, the same references signals but an increased evaluation window $d = 12$ were used to test the T^2 response to the step fault. For this case, the threshold value was calculated to be $T^2_\alpha = 26.569$ and the number of principal components $a = 13$ were used. From Figure 5.6, the T^2 stays within the threshold value before the step fault starts at $j = 200$, and the T^2 value increases to a large value and violates the threshold after the fault occurs. The first time that the T^2 value exceeds its threshold value is around $j = 211$, indicating that there is a delay in fault detection when the evaluation window. The delay is due to the fact that it takes time for the effect of faults to fully influence the evaluation window. There is a trade-off in sensibility and reliability of fault detection by choosing a suitable length of the evaluation window. Note that the T^2 has a significantly reduced oscillation for $d = 12$ in contrast to the very large oscillation for $d = 4$ in Figure 5.4.

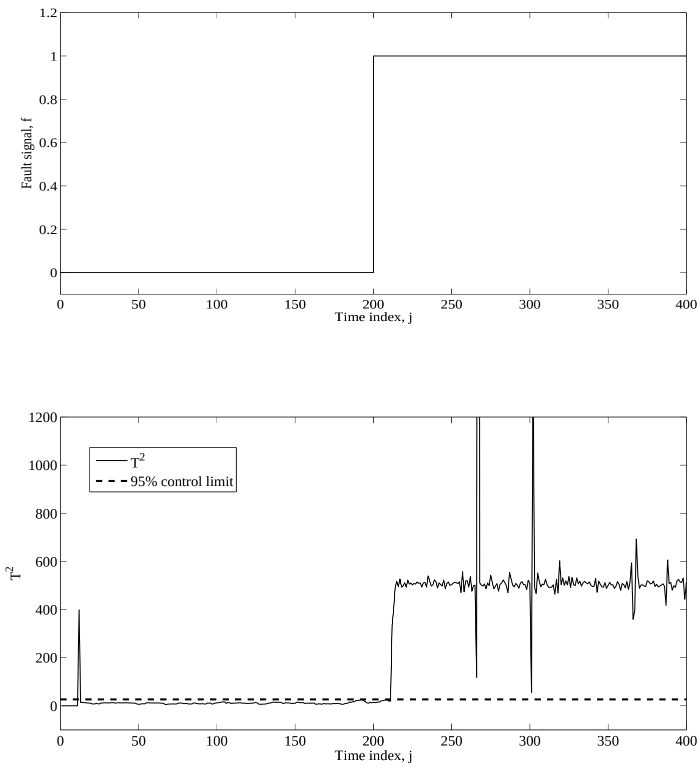


Figure 5.6: Fault detection for a unit step fault with the residual signal generated from the Kalman filter based residual generator and evaluated using DPCA with $d = 12$.

Example 2. In Chapter 3, an observer based fault detection is developed for systems modeled by the Roesser model and a residual signal is obtained from a residual generator. The residual signal is capable of reflecting occurrences of faults. In this example, the obtained residual signal is evaluated using DPCA in order to identify faults clearly. The Example 3 in

Chapter 3 is considered:

$$
\begin{aligned}
\begin{bmatrix} x^h(i+1,j) \\ x^v(i,j+1) \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x^h(i,j) \\ x^v(i,j) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} u(i,j) + \begin{bmatrix} 0 \\ -1 \end{bmatrix} f(i,j), \\
y(i,j) &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x^h(i,j) \\ x^v(i,j) \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \end{bmatrix} d(i,j) + \begin{bmatrix} 0 \\ -1 \end{bmatrix} f(i,j).
\end{aligned} \tag{5.8}
$$

This is a system with an unknown disturbance. The parameters are set as: evaluation window length $d = 6$, the number of principal component $a = 1$ and the number of offline reference database $I = 5$. A unit step fault signal is introduced at $j = 100$. The threshold value was calculated to be $T^2_\alpha = 3.87$. From Figure 5.7, the value of T^2 is zero when there is no fault. At $j = 106$, the T^2 value starts to exceed the threshold value and it increases to a very large value and settles at $T^2 = 2448$. In this case, there are no oscillations in the T^2 value and when the fault occurs, the T^2 value consistently exceeds the threshold value at all time points. This is due to the fact that the system considered contains no noises.

Figure 5.8 shows the T^2 value response if the fault signal is sinusoidal. It is observed that when the fault occurs, the T^2 value increases and exceeds the threshold value. In comparison with that for the step fault in Figure 5.7, the T^2 value is much smaller and just slightly above the threshold value. It would be hard to distinguish the fault by just observing the evolvement of the fault. Using the residual evaluation by DPCA, however, the fault is detected clearly.

5.4 Summary

In model based fault detection for 2-D systems, state estimates are obtained and residual signals are generated using the observer based or Kalman filter based approach. In this chapter, the residual signal is evaluated using DPCA to determine whether a fault occurs with an improved confidence. DPCA has been well accepted for fault detection of 1-D systems. Evaluation of the residual signals of 2-D systems using DPCA provides clear and straightforward indication in fault detection.

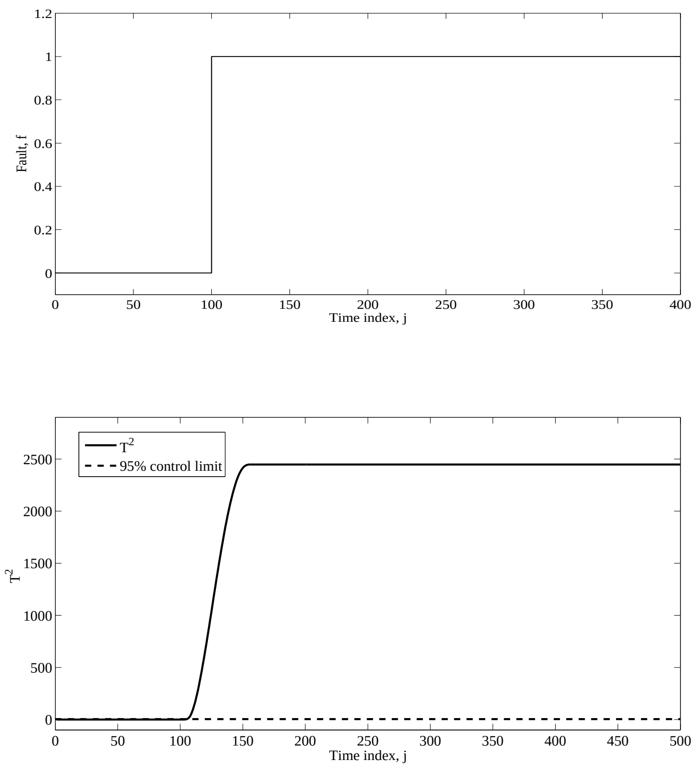


Figure 5.7: Fault detection for a unit step fault with the residual signal generated from the observer based residual generator and evaluated using DPCA.

.

Residual signals reflect the fault occurrences, but when they are corrupted with noises or disturbances, it is reasonable to evaluate them using statistical features and comparing them with threshold values. Based on the historical reference residual databases, the data matrix for the residuals with the previous d observations is formed and the principal components are determined that represent most of variations of the time-lagged matrix. In on-line fault detection, the real time residual signals are calculated and the T^2 value is obtained with the time lagged residual signals. Fault detection can then be performed by comparing the T^2 value with its threshold value. Simulations on two examples were run to evaluate the performance of the fault detection method. Results indicate that the T^2 value stays within the threshold values when there is no fault and exceeds the threshold after a fault occurs. Fault can therefore be clearly detected from evaluation of the residual signal using DPCA. The method is effective for step faults as well

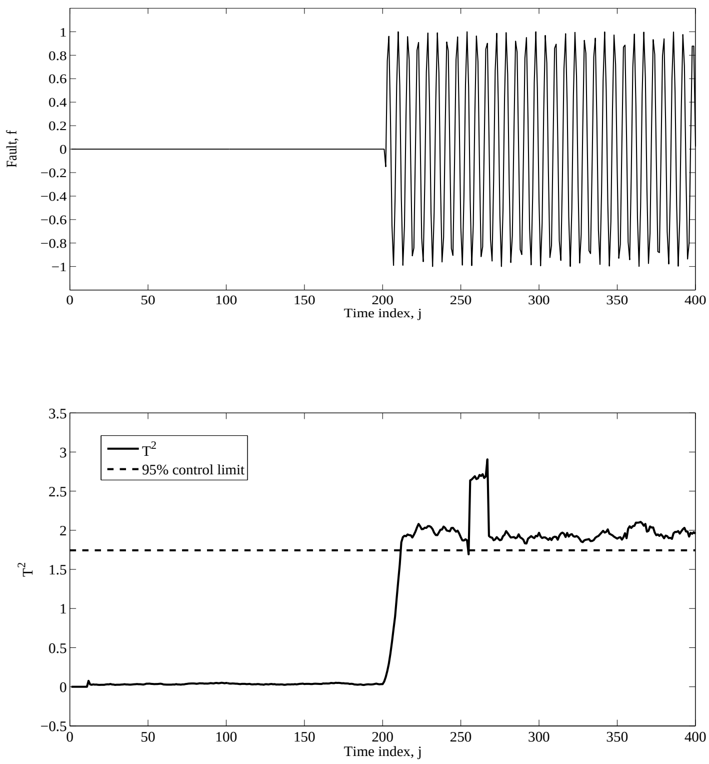


Figure 5.8: Fault detection for a sinusoidal fault with the residual signal generated from the observer based residual generator and evaluated using DPCA.

as faults with zero mean values such as sinusoidal faults. In using DPCA, the current residual signals together with the previous d observations are examined and the length of the evaluation window d affects the performance of the fault detection with reduced oscillations but increased detection delays.

Chapter 6

Conclusions and Future Work

Fault detection is essential in ensuring safe and reliable operations in industrial systems. Extensive research and applications on fault detection of 1-D systems are available. Despite wide existence of 2-D systems, research on fault detection of 2-D systems is few. 2-D systems can generally be described by Roesser model and F-M model. Based on the two model forms, this thesis is focused on developing model based fault detection approached for 2-D systems.

6.1 Conclusions

The observer based fault detection technique has been an active field in the area of control theory and engineering. In Chapter 3, a dead-beat observer based fault detection is developed for Roesser model. The proposed method requires development of an observer and a residual generator. Based on the PBH observability matrix of Roesser model, the transfer function of the observer is obtained by solving a Bézout equation. Application of a realization technique lead to the state space Roesser model of the observer. The state estimate from the observer together with process input and output form the inputs to a residual generator. The residual generator is designed using the polynomial matrices such that the residual directly reflects the fault signal. An efficient realization technique is integrated to the fault detection method to convert the transfer functions of the observer and residual generator to their state space Roesser models. Through the unknown input decoupling, the fault detection approach is able to deal with the systems with unknown inputs or disturbances. Simulation results indicate that the observer based fault detection method can effectively determine whether a fault occurs for systems with or without unknown input. The proposed fault detection method for Roesser model is an extension of the existing work on the dead-beat observer based fault detection for F-M model.

The observer based fault detection method deals with the 2-D systems without noises, but noises are inevitable in real applications. In addition, the observer based fault detection method requires strict conditions that have to be satisfied in order for the observer and residual generator to exist. To overcome the limitations, use of 2-D Kalman filter in fault detection is explored in this thesis. The Kalman filter has been a well accepted method for state estimation and the Kalman filter based fault detection has become a mature technique in 1-D systems. Despite the challenge in extending the well developed 1-D Kalman filter to 2-D systems, the Kalman filters in 2-D domain have been proposed. No research has, however, been reported on fault detection using 2-D Kalman filters. The fault detection using 2-D Kalman filter in this work opens up a new field in 2-D systems.

In Chapter 4, a 2-D Kalman filter is introduced into the FDI framework for the F-M model. A residual signal is generated from the innovation process comparing the output measurement with the state estimate obtained from a recursive 2-D Kalman filter. The generated residual is found to be directly related to fault information. When there is no fault, the residual is zero mean white Gaussian signal. When a fault occurs, the residual is no longer of zero mean white Gaussian process. Due to the noisy nature of systems, it is necessary to evaluate the residual signal over a 2-D evaluation window. The mean of the residual over the evaluated window is formulated describing how the fault affects the residual. Evaluation of the residual is investigated using GLR and a fault is determined by comparing the testing statistic of the residual with its threshold value. The approach is validated with two examples via simulation. For a step fault signal, the residual responds to the faults quickly with non-zero mean and increased variances. If the fault is of zero mean such as a sinusoidal signal, there is no noticeable change in the mean of the residual but significant increase in the variances is observed. The fault can therefore be clearly detected with accuracy using the Kalman filter based approach.

In the fault detection approaches using the observer or the Kalman filter, a residual signal is generated reflecting the fault information. To achieve fault detection with confidence, it is necessary to perform a residual evaluation and determine the fault with a threshold value. In Chapter 4, GLR is applied to the residual evaluation of the Kalman filter based fault detection method. In Chapter 5, an alternative approach is proposed for residual evaluation that is applicable for the residuals from both the observer based and Kalman filter based approaches. The fault detection approach is obtained by applying DPCA to the residual signal obtained from either the observer based or the Kalman filter based residual generator. Based on historical reference data, the time lagged residual data matrices are used to calculate the principal components model of the residual. Online fault detection requires to calculate the T^2 value of the real time residual signal together with its previous d values. Fault is detected with confidence comparing the T^2 value with a threshold value. Examples are simulated for both an F-M model with residual generated from the Kalman filter based approach and a Roesser model with residual from the observer based generator. Results show that a variety of faults can be detected with confidence using the approach.

6.2 Future work

Model based fault detection for 2-D systems is an important but challenging subject. In this thesis, significant progress has been made in developing fault detection methods for 2-D systems. Based on the current work, there are several open problems that need further investigation and may be suggested for future research.

The Kalman filter based fault detection method for F-M model is developed in the thesis. A well-developed Kalman filter for Roesser model has not been available. The challenge lies in the difficulty in calculating the covariances of horizontal and vertical estimation error. Future work is required to develop an efficient Kalman filter based on Roesser model and subsequently a novel Kalman filter based fault detection method.

Model based fault detection depends on availability of models. 2-D model identification techniques are far from being mature. Further research is needed to develop identification techniques for F-M model and Roesser model based on 2-D industrial data. The fault detection methods integrating well developed identification techniques will provide a wider applicability and improved practical implementations.

State estimation is crucial for model based fault detection. The available state estimation methods for 2-D systems, however, have not been abundant. It is,therefore, necessary to explore a variety of state estimation methods in 2-D form including Bayesian estimation, H_2/H_∞ estimation, maximum likelihood estimation and non-linear estimation et al. Enhanced state estimation will bring the fault detection techniques to a higher level.

Bibliography

Antoniou, G.E., Paraskevopoulos, P.N., Varoufakis, S.J. (1988). Minimal state-space realization of factorable 2-D transfer functions. *IEEE transactions on circuits and systems*, 35(8), 1055–1058.

Attasi, S. (1976). Modelling and recursive estimation for double indexed sequences. *Mathematics in Science and Engineering*, 126, 289–348.

Azimi-Sadjadi, M. R., Bannour, S. (1991). Two-dimensional adaptive block Kalman filtering of SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 29(5), 742–753.

Azimi-Sadjadi, M. R., Wong, P. W. (1987). Two-dimensional block Kalman filtering of image restoration. *IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP35(12), 1736–1749.

Basseville, M. (1988). Detecting changes in signals and systems: a survey. *Automatica*, 24 (3), 309–326.

Basseville, M., Benveniste, A. (1986). *Detection of abrupt changes in signals and dynamic systems*. Lecture Notes in Control and Information Sciences, 77, Berlin: Springer-Verlag.

Basseville, M., Nikiforov, I.V. (1993). *Detection of abrupt changes: theory and application*.Information and System Sciences Series, Prentice Hall.

Birgit, J. (2002). *A review on realization theory for infinite-dimensional systems*. PhD Thesis, University Dortmund, Germany.

Bisiacco, M. (1985). On the state reconstruction of 2D systems. *Systems & Control Letters*, 5, 347–353.

Bisiacco, M. (1985). State and Output Feedback Stabilizability of 2D Systems. *IEEE Trans. on Circ. and Sys*, 32, 1246–1254.

Bisiacco, M. (2007). Generalized Luenberger observer for 2D state-space models. *Multidimensional (nD) Systems, 2007 International Workshop on*, Aveiro ,115–122.

Bisiacco, M., Fornasini, E. ,Marchesini, G., (1989). Dynamic regulation of 2D systems: a state-space approach. *Linear algebra and its applications*, 122–124; 195–218.

Bisiacco, M., Valcher M.E. (2004). Unknown input observers for 2D state-space models. *International Journal of Control*, 77(9), 861–876.

Bisiacco, M., Valcher M.E. (2006). Observer-based fault detection and isolation for 2D state-space models. *Multidimensional Systems Signal Process*, 17, 219–242.

Bisiacco, M., Valcher M.E. (2008). Dead-beat estimation problems for 2D behaviors. *Multidimensional Systems and Signal Processing*, 19(3-4), 287–306.

Celebi, M. E., Kurz, L.(1991). A fast recursive two dimensional estimation algorithm. *IEEE international conference of Acoustics, Speech and Signal Processing*, 4, 2945–2948.

Chang, C.T., Hwang, J.I. (1998). Simplification techniques for EKF computations in fault diagnosis suboptimal gains. *Chemical Engineering Science*, 53(22), 3853–3862.

Chen, J., Liu, K. C. (2002). On-line batch process monitoring using dynamic PCA and dynamic PLS Models. *Chemical Engineering Science*, 57, 63–75.

Chen, J., Patton, R.J. (1999). *Robust model-based fault diagnosis for dynamic systems*. Massachusetts: Kluwer Academic Publishers.

Clark, R.N. (1979). The dedicated observer approach to instrument fault detection. In: *Proceedings of the 18th IEEE-CDC*, Ford Lauderdale, Florida, 237–241.

Criftcibasi, Y., Yuksel, O. (1983). Sufficient or necessary conditions for model controllability and observability of Roesser's 2-D system model. *IEEE transactions on automatic control*, AC-27(4), 527–529.

Dai, Y., Zhao, J. (2011). Fault Diagnosis of Batch Chemical Processes Using a Dynamic Time Warping (DTW)-Based Artificial Immune System. *Industrial & Engineering Chemistry Research*, 50, 4534–4544.

Ding, S.(2008). *Model-based fault diagnosis techniques, design schemes,algorithms and tools*. Springer.

Doan, X-T, Srinivasan, R. (2008). Online monitoring of multi-phase batch processes using phase-based multivariate statistical process control. *Computers and Chemical Engineering*, 32, 230-243.

Fathi, Z., Ramirez, W.F., Korbiez, J. (1993). Analytical and knowledge-based redundancy for fault diagnosis in process plants. *AIChE J*, 39, 42–56.

Fornasini, E., Marchesini G. (1976). State-space realization theory of two dimensional filters. *IEEE Trans. Autom. Control*, AC-21(4), 484–491.

Fornasini, E., Marchesini G. (1977). Doubly indexed dynamic systems. *Mathematics system theory*, 12, 59–72.

Fornasini, E., Marchesini, G. (1988). State-space realization theory of two dimensional filters. in *IEEE transactions on automatic control*, AC-21(4), 484–491.

Fornasini, E., Valcher M.E. (1997) nD polynomial matrices with applications to multidimensional signal analysis. *Multidimensional Systems and Signal Processing*, 8, 387–407.

Frank, P. (1996). Analytical and qualitative model-based fault diagnosis, a survey and some new results. *European Journal of Control*, 1, 6–28.

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

Galkowski, K.(2001). *State-space realizations of linear 2-D systems with extensions to the general nD* $(n > 2)$ *case*. Springer.

Gameroa, F.I., Colomer, J., Melendez, J., Warren, P. (2006). Predicting aerodynamic instabilities in a blast furnace. *Engineering Applications of Artificial Intelligence*, 19, 103–111.

Gertler, J. (1991). Analytical redundancy methods in fault detection and isolation. In: *Proceedings of IFAC/IAMCS symposium on safe process* , Baden-Baden.

Gertler, J. (1992). Analytical redundancy methods in fault detection and isolation: survey and synthesis. In: *IFAC symposium on online fault detection and supervision in the chemical process industries*.

Gertler, J. (1993). Residual generation in model-based fault diagnosis. *Control-theory and advanced technology*, 9(1), 259–285.

Gertler, J. (1998). *Fault detection and diagnosis in engineering systems*. Marcel Dekker.

Guiver, J. P., Bose, N. K.(1985). Causal and weakly causal 2-D filters with applications in stabilization. *Multidimensional Systems Theory (N. K. Bose, ed.)*, Dordrecht (NL): D. Reidel Publ. Co., 52–100.

Himmelblau, D.M. (1978). *Fault detection and diagnosis in chemical and petrochemical processes*. Amsterdam: Elsevier press.

Hinamoto, T., Doi, A., Muneyasu, M. (1997). 2-D adaptive state-space filters based on the Fornasini & Marchesini second model. *IEEE Trans. On Circuits And Systems II: Analog And Digital Signal Processing*, 44(8), 667–670.

Hoskins, J.C., Himmelblau, D.M. (1988). Artificial neural network models of knowledge representation in chemical engineering. *Computers and Chemical Engineering*, 12, 881–890.

Isaksson, A. (1993). Analysis of identified 2-D non-causal models. *IEEE transactions on information theory*, 39(2), 525–534.

Jackson, J.E.(1991). *A User's Guide to Principal Components*. Wiley, New York.

Kaczorek, T. (1985). *Two-Dimensional Linear Systems*. Springer-Verlag, Boston.

Kaczorek, T. (2001). Perfect observers for singular 2D linear systems. *Bulletin of the Polish Academy of Sciences, Technical Sciences*, 49(1), 141–147.

Kassidas, A., Taylor, P., MacGregor, J.F. (1998). Off-line diagnosis of deterministic faults in continuous dynamic multivariable processes using speech recognition methods. *Journal of Process Control*, 8, 381–393.

Katayama, T., Kosaka, M. (1979). Recursive Filtering Algorithm for a Two-Dimensional System. *IEEE transactions on automatic control*, AC-23(1), 130–132.

Kaufman, H., Woods, J. W., Dravida, S., Tekalp, A. M. (1983). Estimation and identification of two-dimensional images. *IEEE transactions on automatic control*, AC-28(7), 745–756.

Kawakami, A.(1990). A realization method of two-dimensional rational transfer functions. *IEEE Transactions on Circuits and Systems* , 37 (3), 425–432.

Khosravi, A., Melendez, J., Colomer, J. (2009). Classification of sags gathered in distribution substations based on multiway principal component analysis. *Electric Power Systems Research*, 79, 144–151.

Kourti, T., MacGregor, J.F. (1996). Multivariate SPC methods for process and product monitoring. *Journal of Quality Technology*, 28, 409–428.

Kramer, M. A. (1987). Malfunction diagnosis using quantitative models with non-boolean reasoning in expert systems. *AIChE J.*, 33 (1), 130–140.

Ku, W., Storer, R. H., Georgakis, C. (1995). Disturbance detection and isolation by dynamic principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, 30, 179–196.

Kung, S.Y., Levy, B.C., Morf, M., Kailath, T. (1977). New results in 2-d systems theory, part II: 2-d state-space models realization and the notions of controllability, observability, and minimality. *Proceedings of the IEEE*, 65, 945-961.

Kwan, C., Lewis, F. (1999). A note on Kalman filtering. *IEEE transactions on education*, 42(3), 225–228.

Lee, J.M., Yoo, C.K., Lee, I.B. (2004). Fault detection of batch processes using multiway kernel principal component analysis. *Computer Chemical Engineering*, 28, 1837–1847.

Li, X., Gao, H. (2012). Robust finite frequency H_∞ filtering for uncertain 2-D Roesser systems. *Automatica*, 48, 1163–1170.

Li, X., Gao, H. (2013). Robust finite frequency H_∞ filtering for uncertain 2-D systems: The FM model case. *Automatica*, 49, 2446–2452.

Li, X., Gao, H., Wang, C. (2012). Generalized Kalman-Yakubovich-Popov Lemma for 2-D FM LSS Model. *IEEE Transactions on Automatic Control*, 57(12), 3090–3103.

Makoto, O., Sumihisa, H. (1991). Two dimensional LMS adaptive filters. *IEEE transactions in consumer electronics*, 37(1), 66–73.

Maleki, S., Rapisarda, P., Ntogramatzidis, L., Rogers, E. (2013). A geometric approach to 3D fault identification. In: *Proc. of the 8th International Workshop on Multidimensional Systems nDS'13*, Erlagen, Germany.

Massoumnia, M.A. (1996). A geometric approach to the synthesis of failure detection filters. *IEEE Transactions on Automatic Control*, AC-31, 839–846.

Mehra, R.K., Peschon, J. (1971). An innovations approach to fault detection and diagnosis in dynamic systems. *Automatica*, 7, 637–640.

Mehranbod, N., Soroush, M., Panjapornpon, C. (2005). A method of sensor fault detection and identification. *Journal of Process Control*, 15, 321–339.

Mercere, G., Bako, L., and Lecoeuche, S. (2008). Propagator-based methods for recursive subspace model identification. *Signal Processing*, 88, 468–491.

Morf, M., Levy, B. C.,Kung S. Y., and Kailath, T.(1977). New Results in 2D Systems Theory, Part I and II. *Proc. of IEEE*, 65(6), 861–872; 945–961.

Nam, S.C., Abe, M., and Kawamata, M. (2007). GA-based design of 2D state-space digital filters with linear phase characteristics. *Journal of Circuits, Systems and Computers*, 16(2), 287–303.

Nimmo, I. (1995). Adequately address abnormal situation operations. *Chemical Engineering Progress*, 91(9), 36–45.

Nomikos, P., MacGregor, J.F. (1994). Monitoring batch processes using multiway principal component analysis. *AIChE Journal*, 40, 1361–1375.

Ntogramatzidis, L., Cantoni, M. (2012). Detectability subspaces and observer synthesis for two-dimensional systems. *Multidimensional Systems and Signal Processing: an international journal*, 23, 79–96.

Okada, M. and Sugie, T. (1996). Subspace system identification considering both noise attenuation and use of prior knowledge. In: *Proceedings of the 35th Conference on Decision and Control*, 3662–3667.

Oku, H. and Kimura, H. (1999). A recursive 4SID from the input-output point of view. *Asian Journal of Control*, 1(4), 258–269.

Oku, H. and Kimura, H. (2002). Recursive 4SID algorithms using gradient type subspace tracking. *Automatica*, 38, 1035–1043.

Pearson, R. and Pottmann, M. (2000). Gray-box identification of block-oriented nonlinear models. *Journal of Process Control*, 10, 301–315.

Peretzki, D., Isaksson, A., Bittencourt, A., and Forsman, K. (2011). *Data mining of historic data for process identification*. Technical report no.: LiTH-ISY-R-3039, Linkoping University, Sweden.

Peterka, V. (1981). Bayesian approach to system identification. In *Trends and Progress in System Identification*, Eykhof P. (ed.). Pergamon Press; Oxford, 239–304.

Peternell, K., Cherrer, W., and Deistler, M. (1996). Statistical analysis of novel subspace identification methods. *Signal Processing*, 52, 161–177.

Power, Y., Bahri, P.A. (2004). A two-step supervisory fault diagnosis framework. *Computer Chemical Engineering*, 28, 2131–2140.

Qin, S. J. and Ljung, L. (2003). Closed-loop subspace identification with innovation estimation. In *Proceedings of the 13th IFAC symposium on system identification*, Rotterdam, Netherlands, 887-892.

Qin, S., Lin, W., and Ljung, L. (2005). A novel subspace identification approach with enforced causal models. *Automatica*, 41, 2043–2053.

Qin, S. (2006). An overview of subspace identification. *Computers and Chemical Engineering*, 30, 1502–1513.

Ramos, J. (1993). A subspace algorithm for identifying 2-D separable in denominator filters. *IEEE transactions on circuits and systems II: analog and digital signal processing*, 41(1), 63–67.

Ramos, J., Alenany, A., Shang, H., and Lopes dos Santos, P. (2011). Subspace algorithms for identifying separable in denominator 2-D systems with deterministic inputs. *IET Control Theory and Applications*, 5(15), 1748-1765.

Ramos, J. (1994). A subspace algorithm for identifying 2-D separable in denominator filters. *IEEE Transactions on Circuits and Systems*, 41(1), 63–67.

Rikus, E. (1979). Controllability and Observability of 2-D system. *IEEE transactions on automatic control*, AC-24(1), 121–133.

Rocha P., Willems, J. C. (1991). Controllability of 2-D Systems. *IEEE Trans. Aut. Contr.*, 36(4), 413–423.

Roesser, R.P. (1975). A discrete state space model for linear image processing. *IEEE transactions on automatic control*, AC-20, 1–10.

Rosenbrock, H.H. (1970). *State-space and multivaribale theory*. J. Wiley & Sons, New York.

Ruiz, D., Nougues, J.M., Calderon, Z., Espuna, L., Puigjaner, L. (2000). Neural network based framework for fault diagnosis in batch chemical plants. *Computer Chemical Engineering*, 24, 777–784.

Russell, E.L., Chiang, L.H., Braatz, R. D. (2000). *Data-driven techniques for fault detection and diagnosis in chemical processes*. London: Springer.

Russell, E.L., Chiang, L.H., Braatz, R. D. (2000). Fault detection in industrial processes using canonical variate analysis and dynamic principal component analysis. *Chemometrics and Intelligent Laboratory Systems* 51, 81-93.

Russell, E.L., Chiang, L.H., Braatz, R.D. (2000). *Data-Driven Methods for Fault Detection and Diagnosis in Chemical Processes*. Springer-Verlag, London.

Schuermans, M., Lemmerling, P., and Van Huffel, S. (2003). Structured weighted low rank approximation. *Numerical Linear Algebra*, 10, 1–10.

Schuermans, M., Lemmerling, P., and Van Huffel, S. (2006). Block-row Hankel weighted low rank approximation. *Numerical Linear Algebra*, 13, 293–302.

Schutter, B. De (2000). Minimal state-space realization in linear system theory: An overview. *Journal of Computational and Applied Mathematics*, 121(1-2), 331–354.

Sebek, M.(1992). Polynomial solution of 2-D Kalman-Bucy filtering problems. *IEEE Transactions on Automatic Control*, 37(10), 1530–1533.

Shanks, J., Trietel, S., Justice, J. (1972). Stability and synthesis of two dimensional recursive filters. *IEEE transactions on audio electroacoustics*, AU-20, 115–128.

Sheng, M., Zou, Y. (2007). Kalman filter for generalized 2-D Roesser model. *Journal of China ordnance*, 3(1), 43–48.

Siddiqi, S., Boots, B., and Gordon, G. (2007). A constraint generation approach to learning stable linear dynamical systems. *Advances in Neural Information Processing Systems (NIPS) 21*,1329–1336.

Skogestad, S. and Postlethwaite, I. (1996). *Multivariable Feedback Control*. John Wiley & Sons.

Soderstrom, T. and Stoica, P. (1989). *System identification*. Prentice Hall International. Hertfordshire, UK.

Srinivasan, R., Qian, M. (2005). Offline temporal signal comparison using singular points augmented time warping. *Industrial and Engineering Chemistry Research*, 44 ,4697–4716.

Srinivasan, R., Qian, M. (2006). Online fault diagnosis and state identification during process transitions using dynamic locus analysis. *Chemical Engineering Science* ,61, 6109–6132.

Suresh, B. R., Shenoi, B. A. (1981). New results in two-dimensional Kalman filtering with application to image. *IEEE transactions on Circuit and Systems*, CAS-28(4), 307–319.

Swindlehust, A., Roy, R., Ottersten, B., and Kailath, T. (1995). A subspace fitting method for identification of linear state-space models. *IEEE Transactions on Automatic Control*, 40(2), 311–316.

Teixeira, B. and Aguirre, L. (2011) Using uncertain prior knowledge to improve identified nonlinear dynamic models. *Journal of Process Control*, 21(1), 82–91.

Timmons, W., Chizeck, H., Casas, F., Chankong, V., and Katona, P. (1997). Parameter-constrained adaptive control. *Industrial and Engineering Chemistry Research*, 36, 4894–4905.

Trnka, P. (2007). *Subspace identification methods*. PhD Thesis, Czech Technical University in Prague, Czech Republic.

Trnka P. and Havlena, V. (2009). Subspace like identification incorporating prior information. *Automatica*, 45(4), 1086–1091.

Tulleken, W. (1992). *Grey-box modelling and identification topics*. PhD Dissertation, Delft University of Technology, Delft, The Netherlands.

Van Gestel, T., Suykens, J., Van Dooren, P., and De Moor B. (2001). Identification of stable models in subspace identification by using regularization. *IEEE Transactions on Automatic control*, 46(9), 1416–1420.

Van Overschee, P. and De Moor, B. (1996). *Subspace identification for linear systems, theory, implementation, applications*. Kluwer Academic Publishers.

Vanli, O. and Del Castillo, E. (2007). Closed-loop system identification for small samples with constraints. *Technometrics*, 49, 382–394.

Venkatasubramanian, V., Rengaswamy, R., Yin, K., Kavuri, S.N. (2003). A review of process fault detection and diagnosis Part I , II and III. *Computer & Chemical Engineering*, 27, 293–311; 313–326; 327–346.

Verhaegen, M. (1994). Identification of the deterministic part of MIMO state space models given in innovations form from input-output data. *Automatica*, 30(1), 61-74.

Verhaegen, M. and Verdult, V. (2007). *Filtering and system identificaion: a least squares approach*. Cambridge University Press.

Wahlberg, B., Jansson, M., Matsko, T., and Molander, M. (2007). Experiences from subspace system identification - comments from process industry users and researchers. *Modeling, Estimation, and Control, Lecture Notes in Information and Control Sciences*, Springer, 315–327.

Wang, D. (1998). *Identification and approximation of 1-D and 2-D digital filters*. PhD Thesis, Florida Atlantic University, United states.

Wang, Z., Zhao, J., Shang, H. (2012). A hybrid fault diagnosis strategy for chemical process startups. *Journal of Process Control*, 22, 1287–1297.

Wang, Z., Shang, H. (2014). Observer based fault detection for two dimensional systems described by Roesser models. *Multidimensional Systems and Signal Processing*, DOI: 10.1007/s11045-014-0279-2.

Willsky, A.S., Jones, H.L. (1976). A generalized likelihood ratio approach to the detection and estimation of jumps in linear systems. *IEEE Transactions on Automatic Control*, AC-21, 108–112.

Wise, B.M., Gallagher, N.B. (1996). The process chemometrics approach to process monitoring and fault detection. *Journal of Process Control*, 6(6), 329–348.

Wise, B.M., Gallagher, N.B., Watts, S., White Jr., D.D., Barna, G.G. (1999). A comparison of PCA, multiway PCA, trilinear decomposition and parallel factor analysis for fault detection in a semiconductor etch process. *Journal of Chemometrics*, 13, 379–396.

Woods, J. (1979). Correction on "Kalman filtering in two dimensions". *IEEE transactions on information theory*, IT-25(5), 628–629.

Woods, J., Ingle, V. K. (1981). Kalman filtering in two dimensions: further results. *IEEE transactions on Acoustics, Speech and Signal Processing*, ASSP-29(2), 188–197.

Woods, J., Radewan, C. (1977). Kalman filtering in two dimensions. *IEEE transactions on information theory*, IT-23(4), 473–482.

Wu, L., Ho, D.W.C. (2009). Fuzzy filter design for Ito stochastic systems with application to sensor fault detection. *IEEE Transactions on Fuzzy Systems*, 17, 233–242.

Wu, L., Shi, P., Gao, H., Wang, C. (2008). H_∞ filtering for 2D Markovian jump systems. *Automatica*, 44, 1849–1858.

Wu, L., Su, X., Shi, P. (2011). Mixed H_2/H_∞ approach to fault detection of discrete linear repetitive processes. *Journal of The Franklin Institute*, 348, 393–414.

Wu, L., Yao, X., Zheng, W. (2012). Generalized H_2 fault detection for two-dimensional Markovian jump systems. *Automatica*, 48, 1741–1750.

Xiao, C., Sreeram, V., Liu, W., Venetsanopolos, A. (1998). Identification and model reduction of 2-D systems via the extended impulse response gramians. *Automatica*, 34(1), 93–101.

Xu, H., Lin, Z., Makur, A. (2012). The existence and design of functional observers for two-dimensional systems. *Systems & Control Letters*, 61(2), 362–368.

Xu, H., Zou, Y. (2011). H_∞ control for 2-D singular delayed systems. *International Journal of Systems Science*, 42(4), 609–619.

Xu, H., Zou, Y. (2012). Robust H_∞ filtering for uncertain two-dimensional discrete systems with state-varying delays. *International Journal of Control Automation and Systems*, 8(4), 720–726.

Xu, L., Fan, H., Lin, Z., Bose, N. K. (2008). A direct-construction approach to multidimensional realization and LFR uncertainty modeling. *Multidimensional Systems and Signal Processing*, 19, 323–359.

Xu, L., Wu, L.K., Lin, Z.P., Xiao, Y.G. (2005). On realization of 2D discrete systems by Fornasini-Marchesini model. *International journal of control, Automation and systems*, 3(4), 631–639.

Yang, R., Ntogramatzidis, L., Cantoni, M. (2009). On Kalman filtering for 2-D Fornasini-Marchesini models. *Multidimensional (nD) Systems International Workshop on IEEE*, Thessaloniki, 1–8

Yang, R., Xie, L., Zhang, C. (2006). H_2 and mixed H_2/H_∞ control and two dimensional systems in Roesser model. *Automatica*, 42(9), 1507–1514.

Yao, X., Wu, L., Zheng, W.X., Wang, C. (2011). Fault detection filter design for Markovian jump singular systems with intermittent measurements. *IEEE Transactions on Signal Processing*, 59, 3099–3109.

Ye, S., Wang, W., Zou, Y., Xu, H. (2011). Non-Fragile robust guaranteed cost control of 2-D discrete uncertain systems described by the general models. *Circuits, Systems and Signal Processing*, 30(5), 899–914.

Yoon, S., MacGregor, J.F. (2001). Fault diagnosis with multivariate statistical models part I: using steady state fault signatures. *Journal of Process Control*, 11, 387–400.

Zampieri, S. (1991). 2D residual generation and dead beat observers. *Systems and Control Letters*, 17, 483–492.

Zhang, J. (2006). Improved on-line process fault diagnosis through information fusion in multiple neural networks. *Computer Chemical Engineering*, 30 ,558–571.

Zhang, J. Y., Steenaart, W. (1990). A very fast Kalman filter for image restoration. *IEEE international symposium on Circuit and Systems*, 1, 250–253.

Zhao, P., Yu, D. (1993). An unbiased and computationally efficient LS estimation method for identifying parameters of 2D noncausal SAR models. *IEEE transactions on signal processing*, 41(2), 849–857.

Zhu, Y. (1998). Multivariable process identification for MPC: The asymptotic method and its applications. *Journal of Process Control*, 8(2), 101–115.

Zou, Y., Sheng, M., Zhong, N.F., Xu, S. (2004). A generalized Kalman filter for 2D discrete systems. *Circuits systems signal processing*, 23(5), 351–364.