Fault diagnosis and fault tolerant control of complex process systems

FAULT DIAGNOSIS AND FAULT TOLERANT CONTROL OF COMPLEX PROCESS SYSTEMS

BY

HADI SHAHNAZARI, B.Sc., M.Sc.

A THESIS

SUBMITTED TO THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Doctor of Philosophy (2017) McMaster University

(Chemical Engineering) Hamilton, Ontario, Canada

TITLE:	Fault diagnosis and fault tolerant control of complex process systems
AUTHOR:	Hadi Shahnazari B.Sc., (Technical Inspection Engineering) Petroleum University of Technology, Abadan, Iran M.Sc., (Chemical Engineering) Sharif University of Technology, Tehran, Iran
SUPERVISOR:	Dr. Prashant Mhaskar
NUMBER OF PAGES:	xxiii, 190

To my dear parents

Lay Abstract

Automation is the key to increase efficiency and profitability of the processes. However, as the level of automation increases, major control equipment are more prone to faults. Thus, fault detection and isolation (FDI) and fault tolerant control (FTC) frameworks are required for fault handling. Fault handling, however, can only be efficiently achieved if the designed FDI and FTC frameworks are able to deal with complexities arising in process systems such as nonlinearity, uncertainty, high dimensionality and the resulting effects of the existence of complexity in system structure such as faults that cannot be isolated.

This motivates design of FDI and FTC frameworks for complex process systems. First, FDI frameworks are presented that can diagnose faults in the presence of complexities mentioned above. Then, an integrated framework is designed for diagnosing and handling faults of heating, ventilation and air conditioning (HVAC) systems as an industrial case study of complex process systems.

Abstract

Automatic control techniques have been widely employed in industry to increase efficiency and profitability of the processes. However, reliability on automation increases the susceptibility of the system to faults in major control equipment such as actuators and sensors. This realization has motivated design of frameworks for fault detection and isolation (FDI) and fault tolerant control (FTC). The success of these FDI and FTC mechanisms is contingent on their ability to handle complexities associated with process systems such as nonlinearity, uncertainty, high dimensionality and the resulting effects of the existence of complexity in system structure such as faults that cannot be isolated. Motivated by the above considerations, this thesis considers the problem of fault diagnosis and fault tolerant control for complex process systems.

First, an FDI framework is designed that can detect and confine possible locations for faults that cannot be isolated. Next, the problem of simultaneous actuator and sensor fault diagnosis for nonlinear uncertain systems. The key idea is to design FDI filters in a way they account for the impact of uncertainty explicitly. This work then considers the problem of simultaneous fault diagnosis in nonlinear uncertain networked systems. FDI is achieved using a distributed architecture, comprised of a bank of local FDI (LFDI) schemes that communicate with each other. The efficacy of the proposed FDI methodologies is shown via application to a number of chemical

process examples.

Finally, an integrated framework is proposed for fault diagnosis and fault tolerant control of variable air volume (VAV) boxes, a common component of heating, ventilation and air conditioning (HVAC) systems as an industrial case study of complex systems. The advantages of the proposed framework are diagnosing multiple faults and handling faults in stuck dampers using a safe parking strategy with energy saving capability.

Acknowledgments

I would like to express my gratitude to Dr. Prashant Mhaskar for his support and encouragement during course of this work. Traditionally, everyone points out one of the interesting aspects of working with their supervisor. Mine is the long research rallies (discussions) we had when working on different topics that really was joyful moments for me. Thanks for giving me the opportunity to work on world class research projects.

Last but not least, I would like to thank my parents Tahereh and Hossein for always encouraging me to grow and move forward. You guys are the most inspiring people in my life.

Notation and abbreviations

Notation

n	vector size
$\forall$	for all
$\in$	belongs to
$\|x\|$	the Euclidean norm of vector x
$L_f h$	the Lie derivative of h with respect to the vector field f
x	vector of state variables
$\mathbb{R}^n$	the n-dimensional Euclidean space
u	vector of inputs
y	vector of outputs
$\tilde{y}_i$	vector of outputs shared with the ith subsystem
$\tilde{x}_i$	vector of state variables shared with the ith subsystem
$\hat{x}$	vector of state estimates
t_f	time of fault occurrence
t_d	time of fault detection
u_f	actuator faults
y_f	sensor faults

r	residual
δ	threshold
V	Lyapunov function
Ω	Stability region
Π	Feasibility region

Abbreviations

AF	air flow
AHU	air handling unit
CLF	control Lyapunov function
$CSTR$	continuous-stirred tank reactor
DAT	discharge air temperature
DFO	damper fractional opening
EVO	effective valve opening
FDI	fault detection and isolation
$HVAC$	heating, ventilation and air conditioning
$LFDI$	local fault detection and isolation
MMA	methyl methacrylate
MPC	model predictive control
PCA	principal component analysis
PLS	partial least square

SAT	supply air temperature
SFP	supply fan pressure
SPE	squared prediction error
VAc	vinyl acetate
VAV	variable air volume

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

The last decades have witnessed significant improvements in technology, pushing the design and operation envelope to create more complex dynamical systems. The complexity can be due to nonlinearities, uncertainties, high dimensionality with strong interconnections between subsystems of a network. Aided by the advances in sensing, communicating and computing technologies, operation of complex systems is relying extensively on automated control systems to satisfy simultaneously the (sometimes conflicting) requirements of safety, reliability and profitability. Increased automation, however, also makes the plant susceptible to faults (with incorrect measurements/loss of measurements by sensors and errors or total failures in implementation of the prescribed control action by the actuators being the source of a large number of faults) that can result in substantial financial losses and/or safety hazards if not detected and addressed within a time appropriate to the context of the system dynamics. For instance, the U.S. petrochemical industry loses an estimated 20 $ billion per year

because of abnormalities at oil refineries and chemical plants (see e.g., Nimmo (1995)). The above considerations provide a strong motivation for the development of methods and strategies for the design of novel control and fault-detection and isolation and fault-tolerant control (FTC) algorithms that account for system complexities such as nonlinearity, uncertainty, high dimensionality and the resulting effects of the existence of complexity in system structure such as faults that cannot be isolated. Motivated by the above, the objective of this thesis is to develop a comprehensive and dedicated framework for fault diagnosis and fault handling of complex process systems with an emphasis on industrial applicability of the developed results.

1.2 Background

Fault is an unpermitted deviation of inputs, outputs or process parameters from usual conditions (see e.g., Du (2012)). Base on the fault location, faults are classified in three categories: actuator faults, sensor faults and process faults. Actuator faults can be due to mechanical failures and power losses and result in malfunction of control equipment such as pumps and valves due to mismatch between the prescribed and the implemented control actions (see e.g., Du (2012)). In the presence of faulty actuators, control performance can be jeopardized. Sensor faults are usually caused by discalibration of sensors, degradation of sensing component and short circuits (Du (2012)). If a faulty sensor is utilized for feedback in a control loop, the controller will calculate incorrect control actions. As a result of this, the setpoint cannot be met and the produced material will not have the desired quality. The third type of fault is process fault which includes significant changes in process parameters and large disturbance in the process. Process faults can be due to changes in mode of operation

caused by the other parts of the plant or degradation of the process equipments (see e.g., Du (2012)).

The first step in handling of a fault is fault diagnosis that is defined as determination of kind, size, location and time of occurrence of a fault (see e.g., Blanke *et al.* (2006)). Fault diagnosis includes fault detection, isolation and estimation. Fault detection is determination of the faults present in a system and the time of detection (see e.g., Blanke *et al.* (2006)). Fault isolation is determination of the kind, location and time of detection of a fault and follows fault detection (see e.g., Blanke *et al.* (2006)). Fault detection and isolation is termed FDI. The second step in handling a fault is fault tolerant control (FTC). FTC is controlling the process to achieve desired performance in the presence of faults (see e.g., Blanke *et al.* (2006)) and is usually carried out using a robust control law or modified control law followed by controller reconfiguration.

FDI normally requires a reference model to estimate or predict the state or outputs of a system. A fault is detected when there is discrepancy between the expected values and measurements of a variable. The existing results in the literature on FDI can be classified to causal and non-causal methods based on the type of the model that is being utilized for FDI. The causal models are obtained using the detailed knowledge of the process and its underlying physical principles (see e.g., Afram and Janabi-Sharifi (2014)) or identification techniques (see e.g., Van Overschee and De Moor (2012)). The non-causal models are the models that do not describe the existing causal mechanisms in a system and usually are obtained using statistical methods, artificial intelligence methods, etc. The non-causal FDI approaches have

been successfully applied to industrial system (see e.g., House *et al.* (2001), Mahadevan and Shah (2009) and Tayarani-Bathaie *et al.* (2014)). However, they are mainly able to diagnose simple faults i.e., faults that only affect one variable (see e.g., Yoon and MacGregor (2000)). Causal FDI approaches take advantage of existing analytical redundancies in the process model that provide unique relations between a fault and its symptoms. As a result of this, causal FDI approaches are able to detect and isolate both simple and complex faults (faults that their effect propagates through the system and it affects more than one variable or their effect gets hidden by the controller). Prior to the development of modern system identification techniques, the main shortcoming of causal FDI approaches was their high reliability on first principle models. However, this limitation has been fading due to recent advancements in the area of linear and nonlinear system identification using analytical methods (see e.g., Ljung (1998), Sánchez-Peña *et al.* (2007), Van Overschee and De Moor (2012), Schön *et al.* (2011), Alanqar *et al.* (2015) and Schoukens and Tiels (2017)). Thus, in the rest of this text, the focus is on using causal model based FDI [1] approaches to design FDI frameworks that account for complexities arising in process systems.

The model based methods are based on employing information from the system model to diagnose faults. This approach is based on generating residuals, which are in some sense the difference between the expected and observed process behavior, by utilizing the analytical redundancy provided by the process model to determine expected process behavior. Additionally, thresholds are put in place to account for plant model mismatch and measurement noise with an intent to avoid false alarms. There exists a plethora of results on FDI assuming linear process models dynamics

[1]Note that in the literature, causal and model based FDI are often used interchangeably, as long as it does not cause any ambiguity. The same pattern has been followed in this work.

(see, e.g., Frank (1990), Edwards *et al.* (2000), Venkatasubramanian *et al.* (2003) and Tong *et al.* (2014)). However, these results may not remain effective for nonlinear systems due to strong nonlinear characteristics of the system that behavior of the some of the complex systems exhibit as well.

The FDI problem for nonlinear dynamic models has been considered widely in the literature during the past decade (see, e.g., De Persis and Isidori (2001), Mhaskar *et al.* (2008), Mattei *et al.* (2005), Findeisen *et al.* (2003) and Du and Mhaskar (2014)). Most of the existing results, however, focused on isolation of single actuator or single sensor faults by defining residuals simply as estimation error or its other equivalents (see e.g., Mhaskar *et al.* (2008), Du and Mhaskar (2014)). Recently, results have enabled distinguishing between simultaneous sensor and actuator faults, where a system structure was assumed that enabled detection and isolation of all actuator and sensor faults (see e.g., Du *et al.* (2013)). However, the results in Du *et al.* (2013) are derived while assuming no uncertainty.

Existence of uncertainties is another source of complexity in process systems. The problem of FDI has also been studied for nonlinear systems subject to uncertainty. In Du and Mhaskar (2013), the problem of isolation of complex actuator faults (occurrence of several actuator faults in same order of differentiation) in the presence of uncertainty is handled by explicitly characterizing the way the faults affect the nonlinear process system, and driving the system to a point that enables fault isolation. In Floquet *et al.* (2004), a geometric approach is employed for a class of nonlinear systems to decouple effect of uncertainty and fault (only actuator faults are considered) using sliding mode observers. In Yan and Edwards (2007), under certain matching conditions sliding mode observers are designed to reconstruct the fault signal, while

only considering actuator faults. The problem of FDI in the presence of unstructured uncertainty has also been studied (see e.g., Vemuri and Polycarpou (1997), Zhang *et al.* (2002), Zhang *et al.* (2005), Zhang *et al.* (2010b), Zhang *et al.* (2010a) , Zhang (2011) and Zhang *et al.* (2011)) using adaptive estimation techniques. First, a fault detection scheme is designed which simply uses output estimation error as residual. Then, a bank of fault isolation estimators is designed using adaptive estimation techniques.

Another common property of complex systems is being composed of subsystems that results in the high dimensionality of the system. If the interconnections are weak, a decentralized FDI scheme composed of independent LFDI schemes for each subsystem can be a solution to this problem. In Yan and Edwards (2008), a robust decentralized FDI scheme is designed for actuator fault detection and estimation in large scale systems using sliding mode observers. In Du *et al.* (2011), a robust decentralized FDI scheme is designed for actuator faults diagnosis in network systems. However, there are cases that interconnections between subsystems are not negligible, and would result in poor performance of a decentralized FDI scheme (resulting in false alarms or missed faults). In this case, an alternative solution can be designing an FDI scheme with distributed architecture. In the distributed architecture, a LFDI scheme is designed for each subsystem while the LFDI schemes can communicate to exchange information. The information exchanged can be about recently diagnosed faults in the other subsystems or any other information required by the other LFDI schemes. These realizations have motivated the design of FDI schemes with distributed architecture for networked systems.

In Zhang and Zhang (2012), a distributed actuators FDI scheme is proposed for

a class of interconnected uncertain nonlinear systems using adaptive estimation techniques. In Ferrari *et al.* (2012), a distributed framework is presented for diagnosing single actuator and process faults in nonlinear uncertain large-scale discrete-time systems. In Peng *et al.* (2015), a distributed data based actuator fault identification scheme is presented for linear networked process systems. In Keliris *et al.* (2015), an integrated distributed FD scheme is proposed for detection of sensor and process faults in nonlinear uncertain discrete systems. In Reppa *et al.* (2015), a distributed sensor fault diagnosis for a network of interconnected cyber-physical system (CPS)s presented. In Yin and Liu (2017), a distributed FDI scheme is proposed for cascade networked systems in the absence of uncertainty.

In addition to FDI, there exist a plethora of results in the literature on FTC design. The existing methodologies for FTC are divided in two categories: passive and active methods. Passive fault tolerant control systems refer to the condition where the controller is designed in a way that it is insensitive to a certain restricted set of faults (see e.g., Blanke *et al.* (2006)). Active fault tolerant control systems refer to the case where control reconfiguration is used as fall back plan to recover plant in the faulty situation. In system reconfiguration the faulty components are no longer employed and control law is changed accordingly (see e.g., Blanke *et al.* (2006)).

In Mhaskar *et al.* (2008) and Benosman and Lum (2010) active and passive FTC methodologies are presented for handling actuator faults in nonlinear systems using Lyapunov based controllers, respectively. In Mhaskar (2006) and Zhang *et al.* (2010a), active FTC frameworks are presented for handling actuator faults in nonlinear uncertain systems. For large scale and networked systems, a local FTC scheme is designed for each subsystem to handle actuator faults (see e.g., Peng *et al.* (2015) and Khalili

(2017)). However, these methods are based on the assumption that the process can be continued to operated at the nominal operating point.

In 2008, Gandhi and Mhaskar proposed a safe parking framework for nonlinear systems to handle actuator faults that preclude the possibility of continued operating at the nominal operating point (see e.g., Gandhi and Mhaskar (2008)). The proposed safe-parking framework provides solution for safe operation of the plant under faulty condition until the fault is being repaired and enables effective resumption of the nominal operation upon fault-recovery. In Mahmood *et al.* (2008a), safe-parking problem of nonlinear systems in the presence of uncertainty and lack of measurement has been addressed. In Gandhi and Mhaskar (2009), a safe-parking framework for plant-wide fault tolerant control is presented. In Du and Mhaskar (2011), a safe parking framework has been proposed for switched nonlinear systems with fixed and flexible modes. In Du *et al.* (2012), the existing results for safe parking of nonlinear uncertain systems has been extended to the case where an actuator seizes in an arbitrary value. In Du *et al.* (2011), an integrated framework is presented for fault diagnosis and safe parking of networked systems subject to actuator faults.

In comparison to actuator faults, there exits fewer results on FTC for sensor faults (see e.g., Mhaskar *et al.* (2007) and Du and Mhaskar (2014)). In Du and Mhaskar (2014), a framework for handling sensor faults has been proposed that utilizes the healthy estimates of faulty sensors in the closed loop to maintain nominal operation. For handling sensor faults in nonlinear uncertain systems and networked systems, the proposed methodology in Du and Mhaskar (2014) can be adapted by using the robust estimation filters available in the literatures (see e.g., Tan and Edwards (2003), Yang and Zhu (2015), Yang *et al.* (2015) and Shahnazari and Mhaskar (2016)) and designing

a local FTC scheme for each subsystem, respectively. To handle actuator and sensor faults simultaneously, both sensor and actuator handling approaches would have to be implemented simultaneously (see e.g., Shahnazari *et al.* (2016)). However, the successful handling of simultaneous actuators and sensor faults relies on the ability of FDI schemes to diagnose simultaneous actuator and sensor faults.

1.3 Objectives and outline

A close examination of the literature shows a lack of results for simultaneous actuator and sensor fault diagnosis for a generalized class of nonlinear systems subject to uncertainty in the absence of full state measurements. Also, there is a lack of results in the literature for simultaneous actuator and sensor fault diagnosis in networked systems. In addition, the literature does not provide any insights regarding analysis and classification of faults that cannot be isolated due to lack of existence of enough analytical redundancy in the system structure when it comes to nonlinear systems. Furthermore, while there are some results for isolation of multiple faults and safe parking design applied to simulation case studies (Du and Mhaskar (2014) and Gandhi and Mhaskar (2008)), the literature still lacks from an integrated fault diagnosis and fault handling design with such capabilities that can be applied to industrial case studies. Motivated by the above, the objectives of this thesis are as follows:

1. To explore what are the necessary and sufficient conditions for isolation of faults in nonlinear systems.

2. To develop a framework for simultaneous actuator and sensor fault diagnosis

that explicitly accounts for process nonlinearities, uncertainties and the unavailability of full state measurements.

3. To develop a framework for simultaneous actuator and sensor fault diagnosis in networked system while explicitly accounting for high dimensionality, process nonlinearities, uncertainties, the unavailability of full state measurements and fault in the shared interconnections of the network.

4. To develop an integrated framework for fault diagnosis and fault handling of variable air volume (VAV) of heating, ventilation and air conditioning (HVAC) systems as an industrial case study of complex systems.

The rest of the thesis is organized as follows:

In Chapter 2, the problem of detecting and isolating distinguishable actuator and sensor faults in the solution copolymerization of methyl methacrylate and vinyl acetate monomers is addressed. To this end, first state estimates are generated using a bank of high-gain observers, and nonlinear fault detection and isolation (FDI) residuals are defined. The process dynamics are further analyzed to categorize fault scenarios as distinguishable and indistinguishable, and the necessary and sufficient conditions for the classification are presented. Subsequently, filters are designed that enable FDI for the distinguishable fault scenarios, with the advantage of detecting and confining possible locations for indistinguishable faults. The FDI filters are implemented on the copolymerization process, and the results compared with a linear model based filter design.

In Chapter 3, the problem of simultaneous actuator and sensor fault detection and isolation (FDI) is addressed for control affine nonlinear uncertain systems in the

absence of measurement noise. FDI is achieved by using a bank of filters which utilize a subset of the measurements along with prescribed values of the control actuators to estimate states and compute expected process behavior. Residuals are next defined as the difference between the observed and expected behavior. Detectability conditions are developed which, upon satisfaction, ensure that each residual remains sensitive to a subset of fault scenarios in the presence of uncertainty. To this end, first the ability of observers in providing bounded estimation error for a generalized class of nonlinear uncertain systems is rigorously established. These bounds allow determining thresholds that account for the impact of uncertainty on each residual. Finally, the ability of the proposed framework to achieve fault detection and isolation (FDI) by ensuring a unique residual breaching pattern for each fault scenario is established. The efficacy of the fault detection and isolation framework subject to uncertainty and measurement noise is illustrated using a chemical reactor example.

In Chapter 4, the problem of simultaneous fault diagnosis for nonlinear uncertain networked systems is addressed utilizing a distributed fault detection and isolation (FDI) strategy. The idea is to design a bank of local FDI (LFDI) schemes that communicate with each other. The proposed distributed FDI scheme is shown to be able to handle local faults as well as those that affect more than one subsystem. This is achieved via appropriate adaptation of the LFDI filter based on information exchange with other subsystems and introducing a new concept called detectability index. The detectability and isoability conditions are rigorously derived for the distributed FDI scheme. Effectiveness of the proposed methodology is shown via application to a reactor-separator process subject to uncertainty and measurement noise.

In Chapter 5, an integrated framework for fault detection and isolation (FDI)

and fault tolerant control (FTC) of variable air volume (VAV) boxes, a common component of heating, ventilation and air conditioning (HVAC) systems is presented. To this end, first a statistical model based FDI framework is designed using existing techniques such as principal component analysis (PCA) and joint angle analysis as a benchmark for comparison. Then a novel linear causal model based framework for FDI of multiple actuator and multiple sensor faults is designed and implemented and shown to possess superior FDI capabilities compared to the statistical model based framework. Finally, a safe parking strategy is designed and the ensuing energy savings for the case of stuck dampers demonstrated.

Chapter 2

Fault detection and isolation analysis and design for solution copolymerization of MMA and VAc process

The contributions of this chapter have been published in:

Journal Papers:

Shahnazari, H., Mhaskar, P., *et al.* (2016). Fault detection and isolation analysis and design for solution copolymerization of MMA and VAc process. *AIChE Journal*, **62**(4), 1054–1064.

Refereed Conference Proceedings:

Hadi Shahnazari and Prashant Mhaskar. Detecting and isolating sensor and actuator faults in solution copolymerization of MMA and VAc process. In *American Control*

Conference (ACC), 2015, pages 1617–1622. IEEE, 2015.

2.1 Introduction

Polymerization processes play an important role in chemical industries. The increasing demand for high quality polymers has motivated significant automation to provide the desired quality in the polymer products. However, as the level of automation increases, the process needs to be safeguarded against actuator and sensor faults. If not properly handled, they may cause issues such as off-spec product, plant shutdowns, economic losses, or even safety hazards. Fault-handling, however, can only be efficiently achieved subsequent to fault detection and isolation (FDI). This realization has driven significant effort in the area of fault detection and isolation.

There exists a plethora of results on FDI assuming linear process dynamics (see, e.g., Frank (1990), Edwards *et al.* (2000), Venkatasubramanian *et al.* (2003) and Tong *et al.* (2014)). However, these results may not remain effective for the copolymerization processes owing to the strong nonlinear characteristics of the process.

The FDI problem has also been considered for nonlinear systems subject to actuator and sensor faults, including approaches that utilize data-driven methods and those that generalize the problem to handle hybrid systems (see, e.g., De Persis and Isidori (2001), Doymaz *et al.* (2001a), Doymaz *et al.* (2001b), Mhaskar *et al.* (2008), Zhang *et al.* (2010b), Chilin *et al.* (2010), Hu and El-Farra (2011), Du and Mhaskar (2013) and Du *et al.* (2013)). However, in some cases, as with the copolymerization process under consideration, the system structure does not allow the isolation of all possible fault scenarios.

Motivated by the above considerations, this chapter considers the copolymerization process and presents an FDI mechanism cognizant of the fact that the system structure permits detection and isolation of only a subset of the faults. Then it is

established that the designed scheme is able to detect all possible fault scenarios and confine possible locations for the fault scenarios that cannot be isolated (indistinguishable) to a subset of the possible fault scenarios.

The rest of this chapter is organized as follows: In Section 2.2 the polymerization process is described and a mathematical model for the process is presented. The control objectives for the polymerization reactor are described in 2.3. Then, for the copolymerization process, a nonlinear actuator and sensor fault detection and isolation framework for the fault scenarios that can be isolated (distinguishable faults) is designed with the advantage of recognizing the distinction between distinguishable fault scenarios and indistinguishable fault scenarios in Section 2.4. As a basis of comparison with existing approaches, linear FDI filters are designed by utilizing a linearized model for the process in Section 2.5. The designed linear and nonlinear FDI frameworks are applied to the copolymerization process in Section 2.6. Finally, the results are summarized in Section 2.7.

2.2 Process Description and Model

In this section, the MMA and VAc solution copolymerization process is described, where monomers A (MMA) and B (VAc) are continuously fed to a continuous-stirred tank reactor (CSTR) with initiator (azobisisobutyronitrile, AIBN), solvent (benzene), and chain transfer agent (acetaldehyde). A cooling jacket is equipped to remove the heat of the copolymerization reaction. The mathematical model for this reactor (in the absence of recycle streams and inhibitors) is of the following form (see Du and

Mhaskar (2013) and Congalidis *et al.* (1989)):

$$\begin{aligned}
\dot{C}_j &= \left(\frac{Q_j}{M_j} - \frac{C_j \sum_j Q_j}{\rho}\right)\frac{1}{V} - R_j, \quad j = a, b, i, s, t \\
\dot{T}_R &= (T_0 - T_R)\frac{\sum_j Q_j}{\rho V} + [(-\Delta H_{paa})k_{paa}C_a C_{a\cdot} + (-\Delta H_{pba})k_{pba}C_a C_{b\cdot} \\
&\quad (-\Delta H_{pab})k_{pab}C_b C_{a\cdot} + (-\Delta H_{pbb})k_{pbb}C_b C_{b\cdot}]\frac{1}{\rho c_p} - \frac{UA(T_R - T_c)}{\rho c_p V}
\end{aligned} \tag{2.1}$$

where C_j is the concentration of species j, with subscript a, b, i, s, and t denoting monomer A, monomer B, initiator, solvent, and chain transfer agent, respectively, T_R is the temperature in the reactor, Q_j is the mass flow rate of species j, T_c is the temperature in the cooling jacket, M_j is the molar mass of species j, V is the volume of the reactor, ΔH is the enthalpy of the reaction, ρ and c_p are the density and the heat capacity of the fluid in the reactor, respectively, U is the overall heat transfer

coefficient, A is the heat transfer area of the reactor, and

$$
\begin{aligned}
R_a &= [(k_{paa} + k_{xaa})C_{a\cdot} + (k_{pba} + k_{xba})C_{b\cdot}]C_a \\
R_b &= [(k_{pbb} + k_{xbb})C_{b\cdot} + (k_{pab} + k_{xab})C_{a\cdot}]C_b \\
R_i &= k_i C_i \\
R_s &= (k_{xas}C_{a\cdot} + k_{xbs}C_{b\cdot})C_s \\
R_t &= (k_{xat}C_{a\cdot} + k_{xbt}C_{b\cdot})C_t \\
C_{a\cdot} &= \frac{-l_2 + \sqrt{l_2^2 - 4l_1 l_3}}{2l_1} \\
C_{b\cdot} &= \beta C_{a\cdot} \\
l_1 &= k_{caa} + k_{daa} + 2\beta(k_{cab} + k_{dab}) + \beta^2(k_{cbb} + k_{dbb}) \\
l_2 &= 0 \\
l_3 &= -2k_i C_i \varepsilon \\
\beta &= \frac{(k_{pab} + k_{xab})C_b}{(k_{pba} + k_{xba})C_a}
\end{aligned}
\tag{2.2}
$$

where $C_{a\cdot}$ and $C_{b\cdot}$ denote the total reactor concentrations of free radicals terminating in A and B, respectively. Each of the rate constants follows Arrhenius dependence on temperature. Thus, for instance:

$$k_{paa} = A_{paa} e^{-E_{paa}/RT_R} \tag{2.3}$$

where A_{paa} and E_{paa} are the preexponential constant and activation energy, respectively, and R is the ideal gas constant. The values of the other preexponential constants and activation energies as well as the rest of the process parameters can be found in Table 2.1 (see also Du and Mhaskar (2013) and Congalidis *et al.* (1989)).

2.3 Control Objective

The control objective under normal conditions is to operate the process at the nominal steady state operating point, $C_{a,n} = 2.5 \times 10^{-1}$ kmol/m^3, $C_{b,n} = 5.84$ kmol/m^3, $C_{i,n} = 2.0 \times 10^{-3}$ kmol/m^3, $C_{s,n} = 2.75$ kmol/m^3, $C_{t,n} = 3.7 \times 10^{-1}$ kmol/m^3, and $T_{R,n} = 350.5$ K, where the subscript n refers to the nominal value of states. It is assumed that all the state measurements are available, and the flow rates Q_j, $j = a$, b, i, s, t, and T_c are chosen as manipulated input variables. The inputs are bounded as $0 \leq Q_a \leq 50$ kg/hr, $0 \leq Q_b \leq 120$ kg/hr, $0 \leq Q_i \leq 0.5$ kg/hr, $0 \leq Q_s \leq 100$ kg/hr, $0 \leq Q_t \leq 10$ kg/hr, and $320 \leq T_c \leq 350$ K. The steady state values of the inputs corresponding to the nominal operating point are $Q_a = 18$ kg/h, $Q_b = 90$ kg/h, $Q_i = 0.18$ kg/h, $Q_s = 36$ kg/h, $Q_t = 2.7$ kg/h, and $T_c = 336.15$ K.

The Lyapunov based nonlinear model predictive control design of Mahmood and Mhaskar (2008) is implemented on the process. The key feature of the MPC design is the implementation of a Lyapunov function decay constraint to achieve stabilization (the formulation is reproduced in the Appendix). The hold-time for the control action is chosen as $\Delta = 6$ min, control horizon $T_c = 2\Delta$, and the prediction horizon $T_p = 10\Delta$. In the objective function, the states are normalized against ranges [0, 1], [0, 8], [0, 5×10^{-3}], [0, 10], [0, 1], and [340, 355], respectively, and the inputs are normalized using the magnitude of constraints. The matrices used to penalize the deviations of the normalized states from the steady state values and the increments of the inputs are chosen as Q_w and R_w, respectively. Q_w is a diagonal matrix with all diagonal arrays equal to 1 and R_w is a diagonal matrix with diagonal arrays equal to 1, 1, 50, 0.5, 1, 1. The Lyapunov function is chosen to be a quadratic function of the form $V(x) = x'Px$, with

$$P = \begin{bmatrix} 22.9 & 3.60 & 3.99\times10^{3} & 0.01 & 5\times10^{-3} & 2.08 \\ 3.60 & 3.41 & 5.3\times10^{2} & 5\times10^{-3} & 5\times10^{-3} & 0.28 \\ 3.99\times10^{3} & 5.3\times10^{2} & 7.98\times10^{5} & 1.24 & 0.28 & 4.49\times10^{2} \\ 0.01 & 5\times10^{-3} & 1.24 & 2.98 & 2\times10^{-3} & 3\times10^{-4} \\ 5\times10^{-3} & 5\times10^{-3} & 0.28 & 2\times10^{-3} & 2.97 & 10^{-4} \\ 2.08 & 0.28 & 4.49\times10^{2} & 3\times10^{-4} & 10^{-4} & 0.52 \end{bmatrix}$$

where the matrix P is obtained by solving the Riccati equation for the linearized system. Note that with the Lyapunov constraint implemented, the other parameters in the MPC can be chosen to reflect relative importance of variables, scale, or other considerations without facing instability. In the present example, the values of parameters Q_w, R_w, T_c and T_p were chosen to achieve reasonable control performance.

Table 2.1: Process parameters for the solution copolymerization example.

Parameter	Value	Unit	Parameter	Value	Unit
V	1	m^3	A_{xba}	5.257×10^4	m^3/kmol·s
R	8.314	kJ/kmol·K	A_{xbb}	1577	m^3/kmol·s
ρ	8.79×10^2	kg/m^3	A_{xbs}	1514	m^3/kmol·s
c_p	2.01	kJ/kg·K	A_{xbt}	4.163×10^5	m^3/kmol·s
U	6.0×10^{-2}	kJ/m^2·s·K	E_i	1.25×10^5	kJ/kmol
A	4.6	m^2	E_{caa}	2.69×10^4	kJ/kmol
T_0	353.15	K	E_{cbb}	4.00×10^3	kJ/kmol
ε	1		E_{daa}	0.0	kJ/kmol
M_a	100.12	kg/kmol	E_{dbb}	0.0	kJ/kmol
M_b	86.09	kg/kmol	E_{paa}	2.42×10^4	kJ/kmol
M_i	164.21	kg/kmol	E_{pab}	2.42×10^4	kJ/kmol
M_s	78.11	kg/kmol	E_{pba}	1.80×10^4	kJ/kmol

M_t	44.05	kg/kmol	E_{pbb}	2.42×10^4	kJ/kmol
A_i	4.5×10^{14}	s^{-1}	E_{xaa}	2.42×10^4	kJ/kmol
A_{caa}	4.209×10^{11}	m^3/kmol·s	E_{xab}	2.42×10^4	kJ/kmol
A_{cbb}	1.61×10^9	m^3/kmol·s	E_{xas}	2.42×10^4	kJ/kmol
A_{daa}	0	m^3/kmol·s	E_{xat}	2.42×10^4	kJ/kmol
A_{dbb}	0	m^3/kmol·s	E_{xba}	1.80×10^4	kJ/kmol
A_{paa}	3.207×10^6	m^3/kmol·s	E_{xbb}	1.80×10^4	kJ/kmol
A_{pab}	1.233×10^5	m^3/kmol·s	E_{xbs}	1.80×10^4	kJ/kmol
A_{pba}	2.103×10^8	m^3/kmol·s	E_{xbt}	2.42×10^4	kJ/kmol
A_{pbb}	6.308×10^6	m^3/kmol·s	$-\Delta H_{paa}$	54.0×10^3	kJ/kmol
A_{xaa}	32.08	m^3/kmol·s	$-\Delta H_{pba}$	54.0×10^3	kJ/kmol
A_{xab}	1.234	m^3/kmol·s	$-\Delta H_{pab}$	86.0×10^3	kJ/kmol
A_{xas}	86.6	m^3/kmol·s	$-\Delta H_{pbb}$	86.0×10^3	kJ/kmol
A_{xat}	2085.0	m^3/kmol·s			

2.4 Fault detection and isolation framework for distinguishable faults in the copolymerization process

In this section, we design a nonlinear actuator and sensor fault detection and isolation framework for the distinguishable faults in copolymerization process. Also, we show the designed framework is able to detect and confine possible locations for indistinguishable faults. For comparison, we also design linear model based FDI filters

by considering a linearized model of the process dynamics.

2.4.1 Nonlinear actuator and sensor fault detection and isolation framework for distinguishable faults in the copolymerization process

The key idea is to exploit the analytical redundancy in the copolymerization process to compute the expected process behavior (see e.g., Du *et al.* (2013)). To achieve this, first state estimates are generated using a bank of high-gain observers Du and Mhaskar (2014). To this end, consider the description of the copolymerization process in the following form:

$$\begin{aligned} \dot{x} &= f(x) + G(x)(u + u_f) \\ y &= h(x) + y_f \end{aligned} \tag{2.4}$$

where $x \in \mathcal{X} \subset \mathbb{R}^n$ denotes the vector of state variables, with $\mathcal{X}$ being a compact set of the admissible state values, $u = [u_1, \ldots, u_m]^{\mathrm{T}} \in \mathbb{R}^m$ denotes the vector of prescribed control inputs, taking values in a nonempty compact convex set $\mathcal{U} \subseteq \mathbb{R}^m$, $u_f = [u_{f_1}, \ldots, u_{f_m}]^{\mathrm{T}} \in \mathbb{R}^m$ denotes the unknown fault vector for the actuators, $y = [y_1, \ldots, y_p]^{\mathrm{T}} \in \mathbb{R}^p$ denotes the vector of output variables, $y_f = [y_{f_1}, \ldots, y_{f_p}]^{\mathrm{T}} \in \mathbb{R}^p$ denotes the unknown fault vector for the sensors, and $G(x) = [g_1(x), \ldots, g_m(x)]$. Due to the presence of physical constraints, the actual input $u + u_f$ implemented to the system takes values from the set $\mathcal{U}$ as well.

The design of the high gain observer requires the satisfaction of Assumption 2.1 below (and exploits the fact that the control action is computed using MPC and held constant over a sampling time):

Assumption 2.1. Findeisen *et al.* (2003) There exist integers ω_i, $i = 1, \ldots, p$, with $\sum_{i=1}^{p} \omega_i = n$, and a coordinate transformation $\zeta = T(x, u)$ such that if $u = \bar{u}$, where $\bar{u} \in \mathcal{U}$ is a constant vector, then the representation of the system of Eq. 2.4 in the ζ coordinate takes the following form:

$$\begin{aligned} \dot{\zeta} &= A\zeta + B\phi(\zeta, \bar{u}) \\ y &= C\zeta \end{aligned} \tag{2.5}$$

where $\zeta = [\zeta_1, \ldots, \zeta_p]^{\mathrm{T}} \in \mathbb{R}^n$, $A = \text{blockdiag}[A_1, \ldots, A_p]$, $B = \text{blockdiag}[B_1, \ldots, B_p]$, $C = \text{blockdiag}[C_1, \ldots, C_p]$, $\phi = [\phi_1, \ldots, \phi_p]^{\mathrm{T}}$, $\zeta_i = [\zeta_{i,1}, \ldots, \zeta_{i,\omega_i}]^{\mathrm{T}}$, $A_i = \begin{bmatrix} 0 & I_{\omega_i - 1} \\ 0 & 0 \end{bmatrix}$, with $I_{\omega_i - 1}$ being a $(\omega_i - 1) \times (\omega_i - 1)$ identity matrix, $B_i = [0^{\mathrm{T}}_{\omega_i - 1}, 1]^{\mathrm{T}}$, with $0_{\omega_i - 1}$ being a vector of zeros of dimension $\omega_i - 1$, $C_i = [1, 0^{\mathrm{T}}_{\omega_i - 1}]$, and $\phi_i(x, \bar{u}) = \phi_{i,\omega_i}(x, \bar{u})$, with $\phi_{i,\omega_i}(x, \bar{u})$ defined through the successive differentiation of $h_i(x)$: $\phi_{i,1}(x, \bar{u}) = h_i(x)$ and $\phi_{i,j}(x, \bar{u}) = \frac{\partial \phi_{i,j-1}}{\partial x}[f(x) + g(x)\bar{u}], j = 2, \ldots, \omega_i$. Furthermore, $T : \mathbb{R}^n \times \mathcal{U} \to \mathbb{R}^n$ and $T^{-1} : \mathbb{R}^n \times \mathcal{U} \to \mathbb{R}^n$ are $\mathcal{C}^1$ functions on their domains of definition.

Assumption 2.1 describes the condition that the nonlinear system of Eq. 2.4 is observable from a given set of measured outputs. The bank of high gain observers is designed by leaving out subsets of the measured variables, subject to the satisfaction of Assumption 2.1 for the remaining measured variables (i.e., verifying whether the states can be estimated using the remaining measured variables). Assumption 2.1 does not hold when either C_s or C_t (or both) are not measured. Note that since C_s and C_t do not appear on the right hand side of any state derivative except $\dot{C}_s$ and $\dot{C}_t$, respectively, they are not observable unless directly measured. Thus the transformation required in Assumption 2.1 only holds for subsets of sensors that

includes both C_s and C_t.

For a particular acceptable choice of a subset of sensors, y, for $t \in [t_k, t_{k+1})$, where $t_k = k\Delta$, $k = 0, \ldots, \infty$, the observer is formulated as follows:

$$\begin{aligned} \dot{\hat{\zeta}} &= A\hat{\zeta} + B\phi_0(\hat{\zeta}, u(t_k)) + H(y - C\hat{\zeta}) \\ \hat{\zeta}(t_k) &= T(\hat{x}(t_k), u(t_k)) \end{aligned} \tag{2.6}$$

where $\hat{x}$ and $\hat{\zeta}$ denote the estimates of x and ζ, respectively, $H = \text{blockdiag}[H_1, \ldots, H_p]$ is the observer gain, $H_i = \left[\frac{a_{i,1}}{\varepsilon}, \ldots, \frac{a_{i,\omega_i}}{\varepsilon^{\omega_i}}\right]^{\mathrm{T}}$, with $s^{\omega_i} + a_{i,1}s^{\omega_i - 1} + \cdots + a_{i,\omega_i} = 0$ being a Hurwitz polynomial and ε being a positive constant to be specified, $\hat{x}(t_k) = T^{-1}(\hat{\zeta}(t_k^-), u(t_{k-1}))$ for $k = 1, \ldots, \infty$, and ϕ_0 is the nominal model of ϕ. The state observer requires the global boundedness of ϕ_0 as it is presented in Assumption 2.2.

Assumption 2.2. Du *et al.* (2013) $\phi_0(\zeta, u)$ is a $\mathcal{C}^0$ function on its domain of definition and globally bounded in x.

In this work, we consider each fault scenario comprising at most two simultaneous faults. Thus with m actuators and p sensors, the total number of possible fault scenarios n_f is

$$n_f = C_1^m C_0^p + C_0^m C_1^p + C_1^m C_1^p + C_2^m C_0^p + C_0^m C_2^p = m + p + mp + \frac{m(m-1)}{2} + \frac{p(p-1)}{2} \tag{2.7}$$

where C_k^n presents the binomial coefficients which is equal to $\binom{n}{k} = \frac{n}{k!(n-k)!}$. For the copolymerization process, given that there are six actuators and six sensors, and considering at most two simultaneous faults, there exist a total $2 \times C_0^6 C_1^6 + C_1^6 C_1^6 + 2 \times C_0^6 C_2^6 = 78$ possible scenarios.

For each fault scenario, the objective is to define a residual as the norm of the

difference between the state prediction and the state estimate for the subsystem (appropriately defined) corresponding to the fault scenario. For a particular fault scenario, the expected process trajectory is computed using the subsystem of the process model that is independent of the specific actuator fault and the state estimates generated by the observer that does not require measurements from that particular faulty sensor, or knowledge of the implemented value of that actuator. This expected trajectory when compared with the observed trajectory to generate residuals results in each residual being sensitive to a unique subset of faults.

Remark 2.1. Note that the proposed FDI scheme in Du *et al.* (2013) does not need any priory knowledge about fault occurrence or location and is based on designing a residual for each fault scenario, i.e., based on a bank of residuals, with unique breaching patterns. Thus, when a particular breaching pattern is observed that matches with a known breaching pattern, that particular fault scenario is deemed to have occurred.

The system structure prerequisite for generating such state estimates is presented in Assumption 2.3 for fault scenarios that are distinguishable. To this end, let $\theta_{f,i}$ denote the fault vector (sensor and/or actuator) for the ith fault scenario with dimension of 1×1 in the case of single fault and 1×2 in the case of a scenario that includes multiple faults, and $\bar{\theta}_{f,i}$ the remaining fault variable vector (the remaining u_f and y_f variables). For example, for a two-input-one-output system (where at most six fault scenarios are possible when under the assumption of no more than two simultaneous faults), $\theta_{f,i}$, $i = 1, \ldots, 6$, can be defined as follows: $\theta_{f,1} = u_{f_1}$, $\theta_{f,2} = u_{f_2}$, $\theta_{f,3} = y_{f_1}$, $\theta_{f,4} = [u_{f_1}, u_{f_2}]^{\mathrm{T}}$, $\theta_{f,5} = [u_{f_1}, y_{f_1}]^{\mathrm{T}}$, and $\theta_{f,6} = [u_{f_2}, y_{f_1}]^{\mathrm{T}}$. The vectors $\bar{\theta}_{f,i}$, $i = 1, \ldots, 6$, can be defined accordingly. For example, $\bar{\theta}_{f,1} = [u_{f_2}, y_{f_1}]^{\mathrm{T}}$ and

$\bar{\theta}_{f,2} = [u_{f_1}, y_{f_1}]^{\mathrm{T}}$. Specifically, let $u_{f,i}$ and $y_{f,i}$ denote the vectors of input and output variables subject to faults $\theta_{f,i}$ in the ith fault scenario. Let $\bar{u}_{f,i}$ and $\bar{y}_{f,i}$ denote the vectors of the rest of the input and output variables in that fault scenario.

Assumption 2.3. Du *et al.* (2013) Consider a particular fault scenario, say the ith. Then assumptions 2.1 and 2.2 hold for the system of Eq. (2.4), with $\bar{u}_{f,i}$ and $\bar{y}_{f,i}$ being the vectors of the rest of the input and output variables, respectively.

Under Assumption 2.3, the jth state observer for the ith fault scenario is designed as follow:

$$\begin{aligned} \dot{\hat{\zeta}}^j &= A^j \hat{\zeta}^j + H^j(\bar{y}_{f,i} - C^j \hat{\zeta}^j) \\ \hat{\zeta}^j(t_k) &= T^j(\hat{x}^j(t_k), \bar{u}_{f,i}(t_k)) \end{aligned} \tag{2.8}$$

where $j = 1, \ldots, p_{ob} + \frac{p_{ob}(p_{ob}-1)}{2}$ where p_{ob} is the total number single system outputs that are observable (i.e., the single system outputs that can be estimated from the rest of measured outputs) and $p_{ob} + \frac{p_{ob}(p_{ob}-1)}{2}$ is the total number of designed observers.

Remark 2.2. Note that in the present manuscript, the high gain observers are only used for the purpose of illustration. Any other estimation scheme such as moving horizon estimation (MHE) could also be used as long as they guaranty fast enough convergence rate. Note that the critical requirement for a successful FDI design is the ability of the state estimates to converge at a sufficiently fast rate, since otherwise the FDI filters will lead to either missed faults or false alarms.

2.4.2 Defining residuals for the copolymerization process

Residual definition for single actuator fault in T_c

We now describe how residuals are generated for the copolymerization process for the fault scenarios for which Assumption 2.3 is satisfied. For example, consider a single actuator fault defined by $\theta_{f,11} = u_{f_6}$ (corresponding to faults in the actuator for T_c), the corresponding state prediction is computed by first considering the subsystem for which $u_6 = T_c$ does not appear on the right-hand side of the corresponding ordinary differential equations (ODE's):

$$\begin{aligned}
\dot{C}_a &= \left(\frac{Q_a}{M_a} - \frac{C_a \sum_j Q_j}{\rho}\right)\frac{1}{V} - R_a(C_a, C_b, C_i, T_R) \\
\dot{C}_b &= \left(\frac{Q_b}{M_b} - \frac{C_b \sum_j Q_j}{\rho}\right)\frac{1}{V} - R_b(C_a, C_b, C_i, T_R) \\
\dot{C}_i &= \left(\frac{Q_i}{M_i} - \frac{C_i \sum_j Q_j}{\rho}\right)\frac{1}{V} - R_i(C_i, T_R) \\
\dot{C}_s &= \left(\frac{Q_s}{M_s} - \frac{C_s \sum_j Q_j}{\rho}\right)\frac{1}{V} - R_s(C_a, C_b, C_i, T_R) \\
\dot{C}_t &= \left(\frac{Q_t}{M_t} - \frac{C_t \sum_j Q_j}{\rho}\right)\frac{1}{V} - R_t(C_a, C_b, C_i, T_R)
\end{aligned} \tag{2.9}$$

The above subsystem is the one that (with appropriate modification) needs to be used in defining the residual. Note that when plant is subject to an actuator fault, the implemented input is the summation of the prescribed input and the faulty input. Thus when an actuator fault takes places, the prescribed input differs from the implemented input. The idea behind the actuator fault isolation is to have a subsystem for which the states continues to match the plant trajectories even when the fault occurs. If this subsystem contained T_R as one of the states, it would have to use T_c (the prescribed value), in the computation which would make the predictions deviate

from the plant behavior. Therefore the subsystem model does not include T_R as one of its states. The prediction model is based on this subsystem, however, wherever T_R appears on the right hand side of corresponding ODE's (which affects the dynamics of the other prediction states), its estimated value, $\bar{T}_R$ is used. Therefore, the model used to generate predictions for this filter takes the following form:

$$\begin{aligned}
\dot{\tilde{C}}_a &= \left(\frac{Q_a}{M_a} - \frac{\tilde{C}_a \sum_j Q_j}{\rho}\right)\frac{1}{V} - R_a(\tilde{C}_a, \tilde{C}_b, \tilde{C}_i, \bar{T}_R) \\
\dot{\tilde{C}}_b &= \left(\frac{Q_b}{M_b} - \frac{\tilde{C}_b \sum_j Q_j}{\rho}\right)\frac{1}{V} - R_b(\tilde{C}_a, \tilde{C}_b, \tilde{C}_i, \bar{T}_R) \\
\dot{\tilde{C}}_i &= \left(\frac{Q_i}{M_i} - \frac{\tilde{C}_i \sum_j Q_j}{\rho}\right)\frac{1}{V} - R_i(\tilde{C}_i, \bar{T}_R) \\
\dot{\tilde{C}}_s &= \left(\frac{Q_s}{M_s} - \frac{\tilde{C}_s \sum_j Q_j}{\rho}\right)\frac{1}{V} - R_s(\tilde{C}_a, \tilde{C}_b, \tilde{C}_i, \bar{T}_R) \\
\dot{\tilde{C}}_t &= \left(\frac{Q_t}{M_t} - \frac{\tilde{C}_t \sum_j Q_j}{\rho}\right)\frac{1}{V} - R_t(\tilde{C}_a, \tilde{C}_b, \tilde{C}_i, \bar{T}_R)
\end{aligned} \tag{2.10}$$

where $\tilde{(\cdot)}$ denotes the predicted value for a particular variable and $\bar{(\cdot)}$ denotes the corresponding estimate.

In Eq. 2.10, the predicted values ($\tilde{C}_j$, where $j = a, b, i, s, t$) are the expected trajectories of states computed using the prediction model presented as Eq. 2.10. The estimate $\bar{T}_R$ for use in the above prediction is generated by designing a high gain observer that uses measurements of $y_1 = C_{\mathrm{a}}$, $y_2 = C_{\mathrm{b}}$, $y_3 = C_{\mathrm{i}}$, $y_4 = C_{\mathrm{s}}$ and $y_5 = C_t$, and computes the state estimates, without requiring knowledge of the true value of T_c. The coordinate transformation for this observer is as follows: $\zeta^4_{1,1} = C_{\mathrm{a}}$, $\zeta^4_{2,1} = C_{\mathrm{b}}$, $\zeta^4_{2,2} = \dot{C}_{\mathrm{b}}$, $\zeta^4_{3,1} = C_{\mathrm{i}}$, $\zeta^4_{4,1} = C_{\mathrm{s}}$, and $\zeta^4_{5,1} = C_{\mathrm{t}}$ where the superscript refers to the observer number, and uses the fact that the input action is computed in a discrete

fashion (see Du and Mhaskar (2014) for further details). The observer design is as follows:

$$\begin{aligned}
\dot{\hat{\zeta}}^4_{1,1} &= \frac{a_{1,1}}{\varepsilon}(y_1 - \hat{\zeta}^4_{1,1}) \\
\dot{\hat{\zeta}}^4_{2,1} &= \hat{\zeta}^4_{2,2} + \frac{a_{2,1}}{\varepsilon}(y_2 - \hat{\zeta}^4_{2,1}) \\
\dot{\hat{\zeta}}^4_{2,2} &= \frac{a_{2,2}}{\varepsilon^2}(y_2 - \hat{\zeta}^4_{2,1}) \\
\dot{\hat{\zeta}}^4_{3,1} &= \frac{a_{3,1}}{\varepsilon}(y_3 - \hat{\zeta}^4_{3,1}) \\
\dot{\hat{\zeta}}^4_{4,1} &= \frac{a_{4,1}}{\varepsilon}(y_4 - \hat{\zeta}^4_{4,1}) \\
\dot{\hat{\zeta}}^4_{5,1} &= \frac{a_{5,1}}{\varepsilon}(y_5 - \hat{\zeta}^4_{5,1})
\end{aligned} \tag{2.11}$$

with $\varepsilon = 0.04$, $a_{i,1} = 5$, and $a_{i,2} = 100$, i =1, 2, 3, 4. The values of $a_{i,1}$, $a_{i,2}$ are selected in a way that they form a Hurwitz polynomial for each subsystem and ε is selected small enough to guaranty fast convergence. The other observers required for the implementation of the rest of the filters are also designed in a similar fashion with the same values of the observer parameters. Among these observers, four are designed using five outputs corresponding to single or multiple fault scenarios (single sensor, single actuators faults, or simultaneous sensor and actuator faults), and six using four outputs, corresponding to multiple fault scenarios (two sensors, or simultaneous sensor and actuator faults).

The dedicated residual for a fault in $u_6 = T_c$ (with $\theta_{f,i} = u_{f_6}$) is then defined as

$$r_{11} = \sqrt{(\tilde{C}_\mathrm{a} - \bar{C}_\mathrm{a})^2 + (\tilde{C}_\mathrm{b} - \bar{C}_\mathrm{b})^2 + (\tilde{C}_\mathrm{i} - \bar{C}_\mathrm{i})^2 + (\tilde{C}_\mathrm{s} - \bar{C}_\mathrm{s})^2 + (\tilde{C}_\mathrm{t} - \bar{C}_\mathrm{t})^2}$$

For each dedicated residual, the breaching pattern can be inferred uniquely. Thus When a fault takes place in T_c only, the estimates of the states (utilized in the present filter) stay accurate, because the prescribed value of T_c is not utilized to generate the

estimates. Furthermore, the subsystem used for prediction has been chosen to be independent of T_c, therefore the predicted values stay the same as the true values, which are in turn being correctly estimated. Thus, this residual stays close to zero. On the other hand, for faults in other actuators and sensors, this residual becomes non-zero. To understand the unique breaching pattern better, consider the residual definition for fault in C_a sensor described next.

Residual definition for single sensor fault in C_a

Consider $\theta_{f,1} = y_{f_1}$, i.e., a scenario where a single fault in $y_1 = C_a$ occurs, for which a residual r_1 needs to be designed. The subsystem appropriate for this filter takes the form:

$$
\begin{aligned}
\dot{\tilde{C}}_a &= \left(\frac{Q_a}{M_a} - \frac{\tilde{C}_a \sum_j Q_j}{\rho}\right)\frac{1}{V} - R_a(\tilde{C}_a, \tilde{C}_b, \tilde{C}_i, \tilde{T}_R)\\
\dot{\tilde{C}}_b &= \left(\frac{Q_b}{M_b} - \frac{\tilde{C}_b \sum_j Q_j}{\rho}\right)\frac{1}{V} - R_b(\tilde{C}_a, \tilde{C}_b, \tilde{C}_i, \tilde{T}_R)\\
\dot{\tilde{C}}_i &= \left(\frac{Q_i}{M_i} - \frac{\tilde{C}_i \sum_j Q_j}{\rho}\right)\frac{1}{V} - R_i(\tilde{C}_i, \tilde{T}_R)\\
\dot{\tilde{C}}_s &= \left(\frac{Q_s}{M_s} - \frac{\tilde{C}_s \sum_j Q_j}{\rho}\right)\frac{1}{V} - R_s(\tilde{C}_a, \tilde{C}_b, \tilde{C}_i, \tilde{T}_R)\\
\dot{\tilde{C}}_t &= \left(\frac{Q_t}{M_t} - \frac{\tilde{C}_t \sum_j Q_j}{\rho}\right)\frac{1}{V} - R_t(\tilde{C}_a, \tilde{C}_b, \tilde{C}_i, \tilde{T}_R)\\
\dot{\tilde{T}}_R &= (T_0 - \tilde{T}_R)\frac{\sum_j Q_j}{\rho V} + [(-\Delta H_{paa})k_{paa}\tilde{C}_a C_{a\cdot} + (-\Delta H_{pba})k_{pba}\tilde{C}_a C_{b\cdot}\\
&\quad (-\Delta H_{pab})k_{pab}\tilde{C}_b C_{a\cdot} + (-\Delta H_{pbb})k_{pbb}\tilde{C}_b C_{b\cdot}]\frac{1}{\rho c_p} - \frac{UA(\tilde{T}_R - T_c)}{\rho c_p V}
\end{aligned}
\tag{2.12}
$$

However, for the purpose of prediction (as before), we need estimates of $\bar{C}_a$ (so we are not forced to use possibly incorrect values of $\bar{C}_a$), a high gain observer is designed that uses measurements of $y_2 = C_\mathrm{b}$, $y_3 = C_\mathrm{i}$, $y_4 = C_\mathrm{s}$, $y_5 = C_t$, $y_6 = T_R$. The coordinate transformation for this observer is as follows: $\zeta^1_{1,1} = C_\mathrm{b}$, $\zeta^1_{1,2} = \dot{C}_\mathrm{b}$, $\zeta^1_{2,1} = C_\mathrm{i}$, $\zeta^1_{3,1} = C_\mathrm{s}$, $\zeta^1_{4,1} = C_\mathrm{t}$, and $\zeta^1_{5,1} = T_\mathrm{R}$ where the superscript refers to the observer number. The observer design is as follows:

$$\begin{aligned}
\dot{\hat{\zeta}}^1_{1,1} &= \hat{\zeta}^1_{1,2} + \frac{a_{1,1}}{\varepsilon}(y_2 - \hat{\zeta}^1_{1,1}) \\
\dot{\hat{\zeta}}^1_{1,2} &= \frac{a_{1,2}}{\varepsilon^2}(y_2 - \hat{\zeta}^1_{1,1}) \\
\dot{\hat{\zeta}}^1_{2,1} &= \frac{a_{2,1}}{\varepsilon}(y_3 - \hat{\zeta}^1_{2,1}) \\
\dot{\hat{\zeta}}^1_{3,1} &= \frac{a_{3,1}}{\varepsilon}(y_4 - \hat{\zeta}^1_{3,1}) \\
\dot{\hat{\zeta}}^1_{4,1} &= \frac{a_{4,1}}{\varepsilon}(y_5 - \hat{\zeta}^1_{4,1}) \\
\dot{\hat{\zeta}}^1_{5,1} &= \frac{a_{5,1}}{\varepsilon}(y_6 - \hat{\zeta}^1_{5,1})
\end{aligned} \tag{2.13}$$

The residual for r_1 is defined as follows:

$$r_1 = \sqrt{(\tilde{C}_\mathrm{a} - \bar{C}_\mathrm{a})^2 + (\tilde{C}_\mathrm{b} - \bar{C}_\mathrm{b})^2 + (\tilde{C}_\mathrm{i} - \bar{C}_\mathrm{i})^2 + (\tilde{C}_\mathrm{s} - \bar{C}_\mathrm{s})^2 + (\tilde{C}_\mathrm{t} - \bar{C}_\mathrm{t})^2 + (\tilde{T}_\mathrm{R} - \bar{T}_\mathrm{R})^2}$$

Following the same lines of arguments as earlier, this residual stays close to zero when the fault scenario involving only a fault in the sensor for C_a takes place.

To recognize that this residual will become non-zero when a fault in say the actuator T_c takes place, first note that the evolution of T_R is directly affected by T_c (and any faults in this actuator). Thus the estimated value of T_R being accurate, it will therefore also be affected by the fault in T_c. However, in the filter, the prediction

model uses the expected or computed value of T_c, therefore the evolution of the states in the prediction model ends up being different from the true evolution. Thus, r_1 becomes non zero and breaches its threshold. The rest of the residuals are designed in a similar fashion, and every filter that uses the computed values of T_c ends up breaching the threshold. In particular, all the residuals breach their thresholds except for r_{11} (the dedicated filter for T_c), r_{12}, r_{13}, r_{14} and r_{15} (filters dedicated for T_c, and a sensor fault). By looking at these residuals, one can deduce that either a fault in T_c, or a fault in T_c and one of the sensors must have taken place. However, since breaching all the dedicated residuals for the sensor faults indicate no fault in any of the sensors, by the process of elimination it can be concluded that a single actuator fault in T_c must have taken place. The rest of the fault detection and isolation logic and uniqueness of the breaching pattern is established in a similar fashion.

Remark 2.3. The system structure does not satisfy the standard assumptions for high gain observer design (e.g., Khalil (1999), Mahmood *et al.* (2008b)). However, the recognition that the control action is implemented in a discrete fashion allows invoking the relaxation on the system structure, as presented in Du and Mhaskar (2014). This in turn enables the design of the high-gain observers required for the purpose of building the bank of observers that constitute the FDI filters.

2.4.3 Indistinguishable faults

Now consider $\theta_{f,i} = y_{f_4}$, i.e., a fault scenario where a fault in $y_4 = C_s$ occurs. Since $\theta_{f,i} = y_{f_4}$ does not include any input fault, the corresponding prediction model takes the same form as it is described in Eq. 2.12. To estimate $\bar{C}_s$, a high gain observer needs to be designed that uses measurements of $y_2 = C_{\mathrm{b}}$, $y_3 = C_{\mathrm{i}}$, $y_4 = C_{\mathrm{s}}$,

$y_5 = C_t$, $y_6 = T_R$. To design the corresponding observer, the required transformation in Assumption 2.1 must exist. However, since C_s is fundamentally unobservable, the required transformation in Assumption 2.1 does not exist and the corresponding high gain observer which is insensitive to fault in $y_4 = C_s$ can not be designed. Note that we are considering the scenario where the sensor for C_s continues to report values. The objective is to design a mechanism to determine whether or not these values are correct. To this end, we need to be able to design residuals in a way mentioned in the manuscript and achieve FDI for sensor faults in $y_4 = C_s$. However, the residual for $\theta_{f,i} = y_{f_4}$ is undefined. When fault in $y_4 = C_s$ takes place, all of the existing residuals (designed for other fault scenarios) breach their thresholds. Similarly, all the residuals breach thresholds when say a fault in $y_5 = C_t$ (another variable that is unobservable) takes place. Thus, just by looking at the residuals, and noting that all of them have breached the threshold, it is not possible to distinguish whether a fault (scenario) has taken place that includes $y_4 = C_s$ or $y_5 = C_t$, or both. Lets consider another fault scenario, denoted by $\theta_{f,i} = u_{f_1}$ with a fault in $u_1 = Q_a$. To define the corresponding prediction model, the subsystem which is not subject to fault input must be used. However, since $u_1 = Q_a$ appears in all of the state equations (see Eq. 2.1), the corresponding prediction model does not exist and as a result the corresponding residual for $\theta_{f,i} = u_{f_1}$ is undefined. Thus also when a fault in $u_1 = Q_1$ takes places, all of the existing residuals breach their thresholds leading to a similar predicament.

In designing the FDI scheme, therefore, it is also important to analyze the possibility of achieving FDI for all possible fault scenarios. To do this rigorously (and

to enable generalization to other systems), we first define distinguishable fault scenarios as those for which if that particular fault scenario occurs, there exists an FDI mechanism that can be used to determine uniquely (based on the evolution of the measurements), the occurrence of that (and only that) fault scenario. Corollary 2.1 presents the necessary and sufficient conditions for a fault scenario to be distinguishable. The proof is omitted here since it follows the same line of arguments as the proof of Proposition 1 and Theorem 1 in Du *et al.* (2013). To this end, let $r_{i,ins}$ denote the vector of corresponding insensitive residuals to the ith fault scenario, as defined in Du *et al.* (2013).

Corollary 2.1. *Consider the system of Eq. 2.4, for which Assumptions 2.1-2.3 hold. A fault scenario $\theta_{f,i}$, where $\theta_{f,i} = y_f$ or $\theta_{f,i} = u_f$ is distinguishable if and only if there exists a one-to-one mapping between every fault scenario and $r_{i,ins}$, where $i \in \{1, \ldots, n_f\}$. Furthermore, any $\theta_{f,i}$ that comprises combinations of distinguishable fault scenarios is distinguishable and combination of an indistinguishable fault scenario with any fault scenario is indistinguishable.*

Corollary 2.1 classifies faults in two categories; distinguishable ($\theta_{f,dis}$) and indistinguishable ($\theta_{f,indis}$) faults. Each fault scenario belongs only to one of these categories. If m_{dis} of the inputs and p_{dis} of the outputs (when considered in isolation) satisfy the conditions in Corollary 2.1, the cardinality of set $\theta_{f,dis}$ is

$$\begin{aligned} C_1^{m_{dis}} C_0^{p_{dis}} + C_0^{m_{dis}} C_1^{p_{dis}} + C_1^{m_{dis}} C_1^{p_{dis}} + C_2^{m_{dis}} C_0^{p_{dis}} + C_0^{m_{dis}} C_2^{p_{dis}} = \\ m_{dis} + p_{dis} + m_{dis} p_{dis} + \frac{m_{dis}(m_{dis}-1)}{2} + \frac{p_{dis}(p_{dis}-1)}{2} \end{aligned} \tag{2.14}$$

Furthermore, since $\theta_{f,dis}$ and $\theta_{f,indis}$ are complement of each other, therefore the cardinality of set $\theta_{f,indis}$ is $n_f - |\theta_{f,dis}|$.

For the copolymerization process $y_4 = C_s$ and $y_5 = C_t$ are fundamentally unobservable, and $u_1 = Q_a$ or $u_2 = Q_b$ or $u_3 = Q_i$ or $u_4 = Q_s$ or $u_5 = Q_t$ appears in all of the state equations resulting in prediction model to be undefined for them. Thus for single fault in any of these actuators or sensors, the corresponding insensitive residuals are undefined. For the copolymerization process, $m_{dis} = 1$, and $p_{dis} = 4$, therefore only $C_0^1 C_1^4 + C_1^1 C_0^4 + C_1^1 C_1^4 + C_2^4 = 15$ scenarios are distinguishable. The rest of the 63 fault scenarios belong to $\theta_{f,indis}$.

With the recognition that some of fault scenarios are indistinguishable, the FDI scheme is still able to detect the indistinguishable faults and confine the possible scenarios for fault location to all possible combinations of indistinguishable actuators and sensors. Theorem 2.1 presents the mechanism for detecting and limiting the possible locations for indistinguishable actuators and sensors. To this end, consider the system of Eq. 2.4, where at most two simultaneous faults can occur and let δ_i denote the threshold for the ith fault scenario.

Theorem 2.1. *Consider the system of Eq. 2.4, for which Assumption 2.1-2.3 hold and $t_{k'}$ be the time (if exists) by which all of the residuals have breached their threshold i.e., $r_{i,k} > \delta_i \ \forall i \in \{1, \ldots, |\theta_{f,dis}|\}$. Consider a time $t_k \geq t_{k'}$ then $\theta_{f,indis}(t) \neq 0$ for some $t \in [t_{k'}, t_k)$.*

Proof. First, note that $r_{i,k} > \delta_i \ \forall i \in \{1, \ldots, |\theta_{f,dis}|\}$, we know that some $\theta_{f_i} > 0$. We then show that some fault scenario $\theta_{f,indis}$ take place by contradiction argument. To this end, lets assume that $\theta_{f,dis}$ take place. Therefore $r_{i,k} \leq \delta_i$ for at least one $i \in \{1, \ldots, |\theta_{f,dis}|\}$ (Theorem 1 in Du *et al.* (2013)). However, this is in contradiction with $r_{i,k} > \delta_i$ for all $i \in \{1, \ldots, |\theta_{f,dis}|\}$. Thus $\theta_{f,indis}(t) \neq 0$ for some $t \in [t_{k'}, t_k)$. This concludes the proof of Theorem 2.1. □

As a result of Theorem 2.1, the design acts as a fault detection mechanism for indistinguishable faults with the ability of confining possible locations for them to a subset of all the fault scenarios. In particular, all of the residuals breaching their thresholds results in detection of an indistinguishable fault. Note that even though precise isolation of the fault is not achieved, the design guarantees determining that one of the indistinguishable sensors and/or actuators i.e., faults in $y_4 = C_s$, $y_5 = C_t$, $u_1 = Q_a$, $u_2 = Q_b$, $u_3 = Q_i$, $u_4 = Q_s$, $u_5 = Q_t$ or any combination of them (including possibly the other sensor and actuators), must have experienced a fault. Note this is a fundamental limitation of the process, and not of the FDI framework.

For instances where it is necessary to isolate all the faults, hardware redundancy i.e., smart sensors can be utilized. Note that the smart sensors and actuators are inherently based on the principle of physical redundancy. For instance, a smart actuator for a valve would have an additional means of 'measuring' the valve opening which can then be compared to the prescribed value to detect and isolate the fault. The proposed FDI approach is not intended to replace the smart sensors and actuators, but instead to complement these, and also point to where such devices are required for the purpose of FDI. In particular, the proposed FDI approach can be utilized, where possible, to achieve FDI for number of sensors and actuators to mitigate the high installation and maintenance cost of smart sensors. Note that a rigorous FDI design points, as with the copolymerization process, to the requirement of smart devices for certain subsets of sensors and actuators where fault isolation is not possible otherwise. Thus, we only use smart sensors for those fault scenarios that can not be isolated using analytical redundancy. This enables us to achieve fault isolation

for each fault scenario, while not requiring smart devices for every sensor and actuator. For more safety critical sensors and actuators, the model based approach can be utilized to provide an additional layer of redundancy to the smart devices.

2.5 Linear FDI filters for the copolymerization process

In this section, we design linear FDI filters for the copolymerization process to compare results obtained from the proposed nonlinear FDI filters. Note that while there exist filters that achieve fault detection (see e.g., Frank (1990), Clark (1978a), Clark (1978b) and Frank (1987)), there is a lack of result in the literature on simultaneous actuator and sensor fault detection and isolation designs. Thus the linear FDI filters are also designed based on Du *et al.* (2013) by considering a linearized model for the copolymerization process. The residuals are defined as norm of difference between prediction model and state estimates in the same fashion as the nonlinear FDI filters. A fault is detected and isolated when the corresponding residuals do not breach their thresholds. As with the nonlinear FDI filters, because of existence of unobservable outputs (C_s and C_t) and appearance of five of the six inputs (Q_a, Q_b, Q_i, Q_s and Q_t) in all of the model equations, we can only design 15 residuals.

2.6 Application of fault detection and isolation framework

In this section, we apply the proposed FDI filters to the process. Practical issues such as parametric uncertainty, time-varying disturbances, and measurement noise are considered in the simulations. Specifically, the values of A_{pbb}, A_{xas}, A_{xbs}, A_{xat}, and A_{xbt} are 10% smaller than their nominal values and A_{xbb} is 10% larger. The bounds on these uncertainty are $\pm 10\%$ of their nominal values. The inlet streams of monomer B and solvent contain small amounts of the other. The mass fraction of monomer B in the flow of solvent varies according to $0.02 + 0.02\sin(t)$, and the mass fraction of solvent in the flow of monomer B varies as $0.01 + 0.01\sin(2t)$. The concentration and temperature measurements have combinations of 5 Hz sinusoidal noises. The measurement noise has a normal distribution of variance 0.02, 0.2, 0.0005, 0.2, 0.02, and 0.5 in C_a, C_b, C_i, C_s, C_t, and T_R, respectively. It is assumed that measurements are sampled 10 times evenly between two successive times when control action is implemented. The noisy measurement are processed through a first order low pass filter with time constant equal to 3 min.

To account for the presence of disturbances and measurement noise, thresholds for each filter were determined based on normal operation, and are reported in Tables 2.2 and 2.3. In particular, the maximum observed value of each residual when operating at steady state, under healthy condition, is selected as the corresponding threshold. It should be noted that threshold values for the linear filters are relatively higher than the nonlinear filters. This is because the estimation error when using the linear model based state observer converges to larger values even in the absence of faults.

Table 2.2: Faults to which the residuals are designed to be insensitive and thresholds for the linear FDI filters.

Residual	Faults	Threshold	Residual	Faults	Threshold
r_1	y_{f_1}	0.3	r_2	y_{f_2}	0.16
r_3	y_{f_3}	0.16	r_4	y_{f_6}	0.17
r_5	y_{f_1}, y_{f_2}	0.16	r_6	y_{f_1}, y_{f_3}	0.16
r_7	y_{f_1}, y_{f_6}	0.002	r_8	y_{f_2}, y_{f_3}	0.16
r_9	y_{f_2}, y_{f_6}	0.002	r_{10}	y_{f_3}, y_{f_6}	0.012
r_{11}	u_{f_6}	0.012	r_{12}	u_{f_6}, y_{f_6}	0.012
r_{13}	u_{f_6}, y_{f_1}	0.012	r_{14}	u_{f_6}, y_{f_2}	0.02
r_{15}	u_{f_6}, y_{f_3}	0.002			

Table 2.3: Faults to which the residuals are designed to be insensitive and thresholds for the nonlinear FDI filters.

Residual	Faults	Threshold	Residual	Faults	Threshold
r_1	y_{f_1}	0.27	r_2	y_{f_2}	0.2
r_3	y_{f_3}	0.07	r_4	y_{f_6}	0.07
r_5	y_{f_1}, y_{f_2}	0.07	r_6	y_{f_1}, y_{f_3}	0.068
r_7	y_{f_1}, y_{f_6}	0.06	r_8	y_{f_2}, y_{f_3}	0.06
r_9	y_{f_2}, y_{f_6}	0.06	r_{10}	y_{f_3}, y_{f_6}	0.06
r_{11}	u_{f_6}	0.01	r_{12}	u_{f_6}, y_{f_6}	0.01
r_{13}	u_{f_6}, y_{f_1}	0.06	r_{14}	u_{f_6}, y_{f_2}	0.01
r_{15}	u_{f_6}, y_{f_3}	0.01			

We first consider two case where a small and a large abrupt, constant bias fault of magnitudes of 0.5 $kmol/m^3$ and 0.05 $kmol/m^3$ in $y_1 = C_a$ (single sensor fault) takes place, respectively, at time $t_f = 1.5$ hr. The evolution of residual profiles is shown in Figure 2.1. It can be seen that some of the filters breach their thresholds for the linear FDI design. In essence, the fault is successfully detected but is not isolated since the residual breaching profiles do not follow any of the expected patterns. For the large abrupt fault, after the fault occurrence, for the first one hour the breaching pattern matches with fault occurrence in $y_1 = C_a$ and after that for one hour and half,

none of the residuals breach their thresholds, (incorrectly) indicating that no fault has occurred. For the small abrupt fault, the breaching pattern does not match with any faulty scenario at any time period. Thus, the linear FDI filters are only able to achieve fault detection regardless of fault magnitude. In contrast, the nonlinear FDI design successfully detects and isolates the fault (see Remark 2.4 for more discussion on this). The evolution of the measurements of the output variables, the state estimates provided by the observer that uses measurements of C_{b}, C_{i}, C_{s}, C_{t} and T_{R}, and the true values of the state variables are depicted by solid, dashed, and dashed-dotted lines in Figure 2.2, respectively.

Remark 2.4. Note that the isolation mechanism is based on some of the residuals being insensitive to each specific fault scenario and some of them not, i.e., based on the existence of a unique breaching pattern for each fault scenario (discussed in Du *et al.* (2013) in more detail). If a residual breaches its threshold, it means a fault has occurred. Fault isolation is achieved by comparing the residuals breaching pattern with patterns of residuals breaching for the different types of faulty scenarios (single sensor, single actuator, two sensor, two actuator, and simultaneous actuator and sensor). For example, in Figure 1 for nonlinear FDI filters, since all of the residuals corresponding to a single sensor except r_1 have breached their thresholds, it can be concluded that a fault in the corresponding sensor to r_1, $y_1 = C_a$ has occurred. Note that r_5, r_6, r_7 and r_{13} also do not breach their thresholds, while the others residuals do. This set of breaching patterns uniquely matches with the breaching pattern for a single fault in the C_a sensor. To understand this better, note that, for instance, r_5 is designed to not breach the threshold when a fault in either the C_a or C_b sensor has taken place. But by also noticing that the dedicated residual for C_b (r_2) has breached

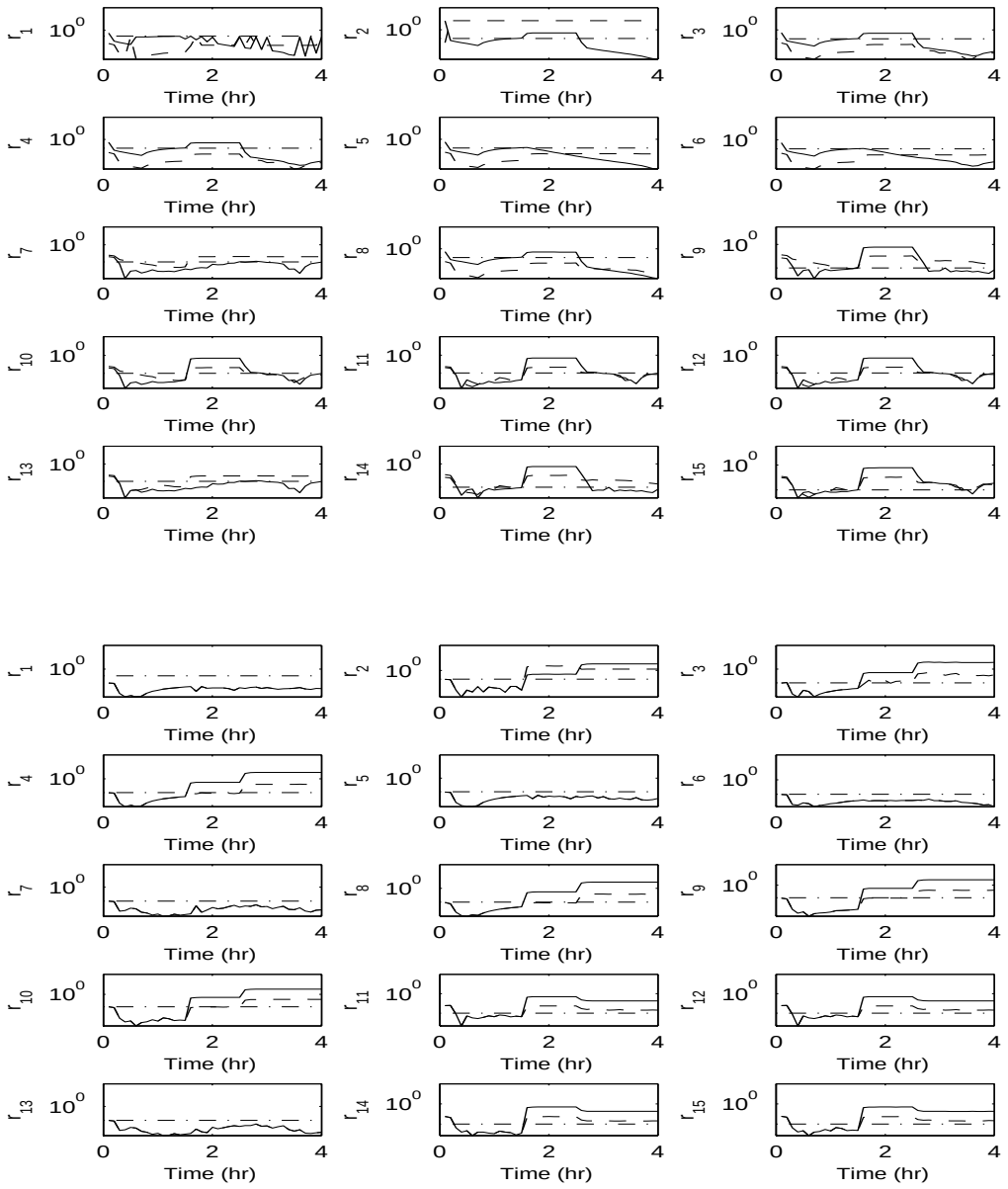


Figure 2.1: Evolution of the residuals for large (solid lines) and small (dashed lines) magnitude constant sensor fault. The thresholds are depicted by the dashed-dotted lines. Top: Using linear FDI filters enables only fault detection for the large and small sensor faults. Bottom: Using nonlinear FDI filters enables FDI for both cases.

the threshold, it is concluded that fault in C_b sensor has not taken place, and thus only a fault in the C_a sensor has occurred (see Du *et al.* (2013) for further discussed along these lines).

We next consider a case where incipient faults in $y_1 = C_{\mathrm{a}}$ and $u_6 = T_{\mathrm{c}}$ (one actuator fault and one sensor fault) take place, starting at time $t_f = 1.5$ hr. The faults are simulated as follows:

$$
\begin{aligned}
y_{f_1} &= \begin{cases} 0, & 0 \le t < t_f \\ (0.05 + 0.05\sin 5t)(2 - e^{5t_f - 5t}), & t \ge t_f \end{cases} \\
u_{f_6} &= \begin{cases} 0, & 0 \le t < t_f \\ (5 + 5\sin 5t)(2 - e^{5t_f - 5t}), & t \ge t_f \end{cases}
\end{aligned}
\tag{2.15}
$$

The evolution of residual profiles is shown in Figure 2.3. Like the previous case, using the linear FDI method, some of the residuals breach their thresholds and therefore the fault is successfully detected. However, the fault is not isolated since residual profiles do not follow any of expected patterns. In contrast, by using the proposed method, the fault in y_1 and u_6 is successfully isolated.

The FDI results for other distinguishable faults scenarios were also considered and are not presented here for sake of brevity. Finally, we demonstrate the ability to detect the indistinguishable faults. In particular, we consider a case where faults

take place in $y_4 = C_s$ and $u_1 = Q_a$ at time $t_f = 1.5$hr, simulated as follows:

$$y_{f_4} = \begin{cases} 0, & 0 \leq t < t_f \\ (0.55 + 0.55\sin 5t)(2 - e^{5t_f - 5t}), & t \geq t_f \end{cases}$$
$$u_{f_1} = \begin{cases} 0, & 0 \leq t < t_f \\ (3.5 + 3.5\sin 5t)(2 - e^{5t_f - 5t}), & t \geq t_f \end{cases} \tag{2.16}$$

The evolution of residual profiles is shown in Figure 2.4. Following Theorem 2.1, since all the residuals breach their thresholds, the fault is successfully detected. However, the fault can not be in any of considered fault scenarios in FDI filters design, since all of the residual have breached their thresholds. Therefore according to the Theorem 2.1, any of the other (indistinguishable) fault scenario must have occurred.

Remark 2.5. In essence, the linear FDI method is only able to detect, but not isolate faults in the copolymerization process. This is primarily due to the estimation and prediction errors associated with using a linear model in the observer, prediction model and filter design. The observer can readily be replaced by other observers (such as the Kalman filter, extended Kalman filter, or the moving horizon observer), to possibly improve the estimation accuracy of the observer; however, the errors associated with prediction using a linear model will still limit the effectiveness of linear model based FDI designs.

Remark 2.6. Note that the presence of the FDI mechanism enables FDI in the closed-loop system thereby allowing the operator to determine the appropriate course of action following a fault. For instance, in the case of a single actuator fault, if the

fault is simply a constant bias fault, then a robust/offset free controller would still keep the process operating at the desired operating point. Knowing through the FDI mechanism that a sensor fault has not taken place can help the operator schedule the correction of such an actuator at a later stage. On the other hand, with the FDI determining a single actuator fault, if the functioning sensors reveal that the process is moving off-spec, or the control action starts chattering (perhaps because the existing robust/offset free controller is not able to handle the particular kind of actuator fault), the operator can then trigger reconfiguration (e.g., Mhaskar *et al.* (2008)) and actuator repair on a more urgent basis. For sensor faults, on the other hand, even if its a constant bias fault, there exists no control law that can drive the process to the desired set-point for the variable in question. The FDI information then becomes critical in taking that sensor out of the loop (where possible), or triggering immediate rectification of the sensor to preserve on-spec production.

Remark 2.7. From fault handling perspective, sensor faults can be handled by using estimation of states that are verified to be accurate, instead of using a faulty sensor reading (as it is done in the simulation results corresponding to Figure 2.2 or see e.g., Du and Mhaskar (2014)). The actuator faults on the other hand directly impact the control action implemented on the plant, and if not handled, could result in the states deviating from nominal operating point. In this case, once such a fault has been detected, robust control methods, or control reconfiguration methods can be used to achieve fault-tolerant control (see, e.g., Mhaskar *et al.* (2008)). To handle actuator and sensor faults simultaneously, both sensor and actuator handling approaches would have to be implemented simultaneously.

2.7 Conclusions

This work considered the problem of isolating distinguishable actuator and sensor faults in the solution copolymerization of MMA and VAc. To achieve fault detection and isolation for the distinguishable faults in copolymerization reactor, an actuator and sensor fault detection and isolation framework was designed. To this end, first state estimates were generated using a bank of high-gain observers and then nonlinear fault detection and isolation (FDI) residuals were defined. The ability of the proposed framework in detecting and narrowing the possible locations for indistinguishable fault scenarios to a subset of possible scenarios was proved and verified through simulations. Illustrative linear FDI filters were also designed for the purpose of comparison. While linear model based FDI only achieved fault detection, the application of the proposed FDI mechanism was found to also successfully isolate distinguishable faults even in the presence of plant-model mismatch and measurement noise.

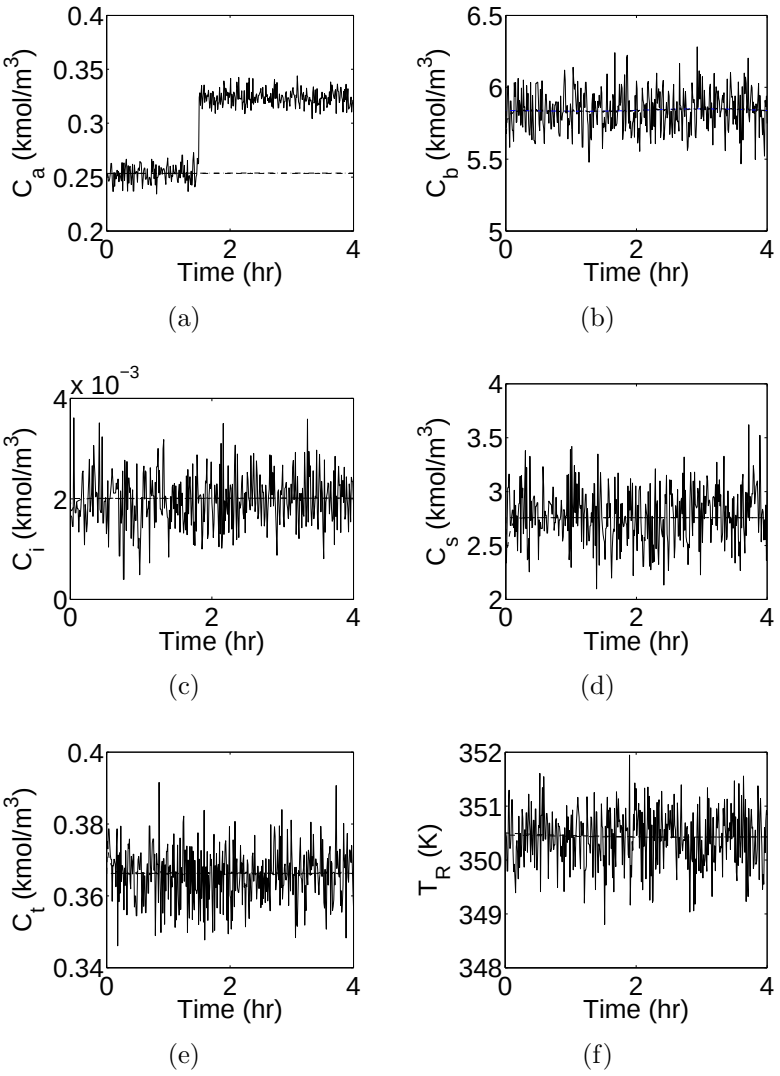


Figure 2.2: Evolution of the closed-loop measurements (solid lines), the state estimates (dashed lines), and the true values of the process states (dashed-dotted lines). A fault takes place in C_a sensor at time $t_r = 1.5$ hr and is handled. Since the observer does not use measurements of C_a, the state estimates stay close to their true values even after the fault takes place.

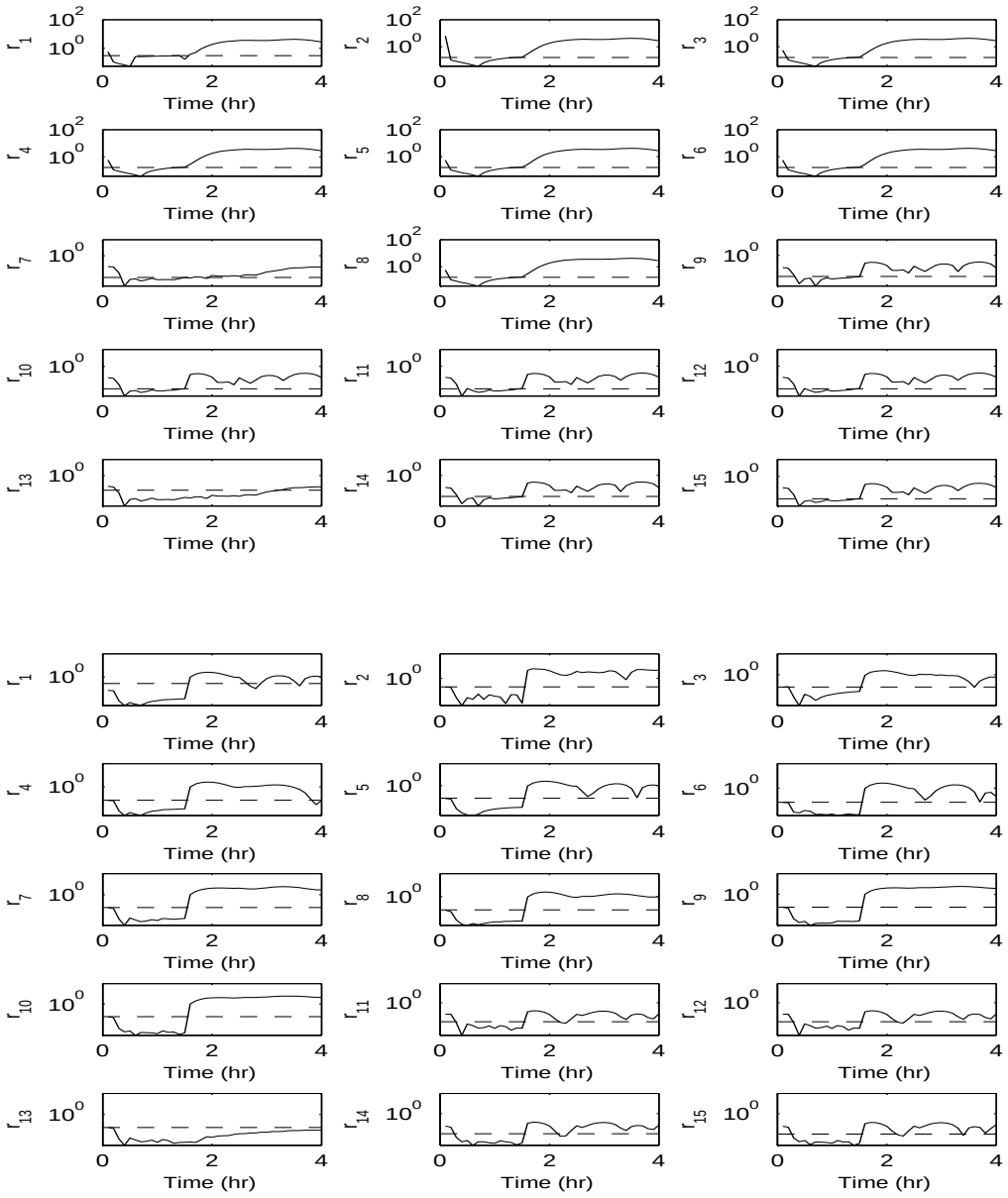


Figure 2.3: Evolution of the residuals (solid lines) and thresholds (dashed lines). Top: Using linear FDI filters: Since all of the residuals breach their thresholds, the fault is detected but is not isolated. Bottom: Using nonlinear FDI filters: Since all the residuals breach their thresholds except for r_{13}, which is insensitive to y_{f_1} and u_{f_6} (see Table 2.3), faults in y_1 and u_6 are isolated.

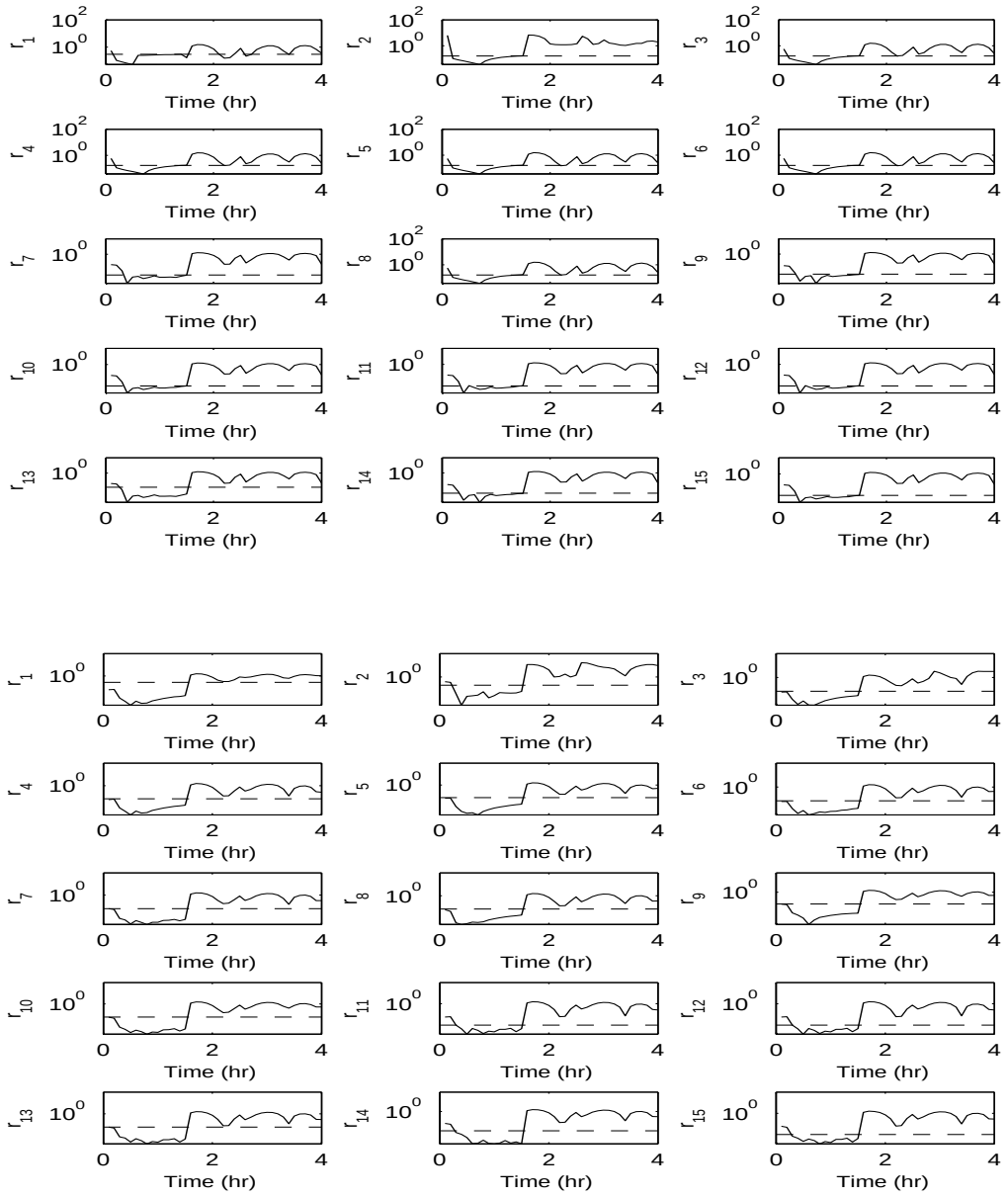


Figure 2.4: Evolution of the residuals (solid lines) and thresholds (dashed lines). Top: Using linear FDI filters, Bottom:Using nonlinear FDI filters. In both cases, since all the residuals breach their thresholds, the fault is detected but is not isolated.

Chapter 3

Actuator and sensor fault detection and isolation for nonlinear systems subject to uncertainty

The contributions of this chapter have been submitted to/published in:

Journal Papers:

Shahnazari, H. and Mhaskar, P. Actuator and sensor fault detection and isolation for nonlinear systems subject to uncertainty, *International Journal of Robust and Nonlinear Control*, Accepted, In press.

Refereed Conference Proceedings:

Shahnazari, H. and Mhaskar, P. (2016). Simultaneous actuator and sensor fault isolation of nonlinear systems subject to uncertainty. In *American Control Conference (ACC), 2016*, pages 6857–6862. IEEE.

3.1 Introduction

Fault detection and isolation (FDI) are critical components of a fault tolerant control system, and are becoming increasingly important given the ubiquitousness of automation- from process industries to self-driven vehicles. The FDI problem explicitly accounting for system nonlinearity has been considered widely in the literature during the past decade, with results often focusing only on actuator or sensor faults (see, e.g., Kaboré *et al.* (2000), De Persis and Isidori (2001), Kabore and Wang (2001), Mhaskar *et al.* (2008), Mattei *et al.* (2005), Ma and Yang (2012), Du *et al.* (2013), Du and Mhaskar (2014), Peng *et al.* (2015) and Shahnazari *et al.* (2016)). In more recent results Du *et al.* (2013) the problem of distinguishing between sensor and actuator faults (albeit in the absence of uncertainty), is addressed. However, the success of FDI mechanisms is contingent on their ability to handle system nonlinearity and uncertainty in a unified framework.

There is a plethora of results in the literature for fault diagnosis of nonlinear uncertain systems (see e.g., Du and Mhaskar (2013), Yan and Edwards (2007), Ma and Yang (2013), Zhu and Yang (2013), Fekih (2014), Zhang *et al.* (2010b), Zhang (2011)). However, the results mentioned in the above only consider single actuator and sensor faults for specific classes of nonlinear uncertain systems. There also exist results in the literature for simultaneous actuator and sensor fault identification for nonlinear uncertain systems (see e.g., Yang and Zhu (2015) and Yang *et al.* (2015)), however they only consider nonlinearities that can be bounded everywhere using the same Lipschitz constant, and the results also only hold for bounded actuator faults. In summary, there is a lack of results for nonlinear systems subject to uncertainty where the problem of fault detection and isolation for simultaneous actuator and sensor

faults is addressed for a general class of fault functionality and system nonlinearity.

Motivated by the above considerations, this chapter considers the problem of actuator and sensor fault detection and isolation for control affine nonlinear systems in the presence of uncertainty. This is achieved by building a bank of residuals, each using an appropriate subset of the available measurements (and associated state observers), to determine the expected behavior of the system and compare with the observed evolution. The residuals are designed to be sensitive to a subset of faults and insensitive to the rest in the absence of uncertainty. To achieve FDI in the presence of uncertainty, thresholds are defined in a way that they account for the impact of the uncertainty on the estimation error and the prediction of the expected system behavior. In this way, each residual is still insensitive to a subset of faults in the presence of uncertainty and sensitive to the rest if the fault functionality satisfies a rigorously derived detectability condition.

The rest of this chapter is organized as follows: the system description is presented in Section 3.2. In Section 3.3, the boundedness of estimation error in the presence of uncertainty using high gain observers is rigorously established. In Section 3.4, the FDI mechanism is presented. In particular, thresholds are defined in a way they utilize the bound determined in Section 3.3 to ensure that there is no false alarms and the residuals retain their property of being insensitive to a subset of faults in the presence of uncertainty. Then the detectability and isolability conditions for single and simultaneous faults are presented, where the detectability analysis establishes that the sensitive property of residuals is also retained in the presence of uncertainty. The efficacy of proposed FDI framework in the presence of uncertainty and measurement noise is illustrated using a chemical reactor example in Section 3.5.

Finally, Section 3.6 presents some concluding remarks.

3.2 Preliminaries

Consider a multi-input multi-output nonlinear system described by

$$\begin{aligned} \dot{x} &= f(x) + G(x)(u + u_f) + \theta(x, u, t) \\ y &= h(x) + y_f \end{aligned} \tag{3.1}$$

where $x \in \mathcal{X} \subset \mathbb{R}^n$ denotes the vector of state variables, with $\mathcal{X}$ being a compact set of the admissible state values, $u = [u_1, \ldots, u_m]^{\mathrm{T}} \in \mathbb{R}^m$ denotes the vector of prescribed control inputs, taking values in a nonempty compact convex set $\mathcal{U} \subseteq \mathbb{R}^m$, $u_f = [u_{f_1}, \ldots, u_{f_m}]^{\mathrm{T}} \in \mathbb{R}^m$ denotes the unknown fault vector for the actuators, θ denotes the uncertainty with $\|\theta(x, u, t)\| \leq \bar{\theta}$, where $\bar{\theta}$ is a known positive constant, $y = [y_1, \ldots, y_p]^{\mathrm{T}} \in \mathbb{R}^p$ denotes the vector of output variables, $y_f = [y_{f_1}, \ldots, y_{f_p}]^{\mathrm{T}} \in \mathbb{R}^p$ denotes the unknown fault vector for the sensors and $G(x) = [g_1(x), \ldots, g_m(x)]$. The inputs are implemented in a discrete fashion, with sampling time Δ. Due to the presence of physical constraints, the actual input $u + u_f$ implemented to the system takes values from the set $\mathcal{U}$ as well. t_a and t_s denote the time of fault occurrence for actuator and sensor faults, respectively. Note that since the main objective of this work is the diagnosis of simultaneous actuator and sensor faults, in the rest of the manuscript, only one time of fault occurrence, t_f, is used. However, the FDI methodology presented in this work is applicable to the cases where actuator and sensor faults do not take place simultaneously. Throughout the manuscript, $L_f h(x)$ denotes the standard Lie derivative of a scalar function $h(x)$ with respect to a vector

function $f(x)$ defined as $L_f h(x) = \frac{\partial f}{\partial x} \cdot h(x)$, and $\| \cdot \|$ denotes the Euclidean norm for the vectors.

Remark 3.1. Note that the above system consideration and uncertainty description accounts for unstructured uncertainty which results from, for instance, fewer number of states, parameters varying with time, state, modeling errors and plant model mismatch. Thus, the developed framework enables fault diagnosis for a generalized class of uncertain systems. Note that while considering parametric (structured) uncertainty addresses a limited class of uncertain systems, there is no guarantee that the resulting bounds on the uncertainty will be tighter with respect to the systems with unstructured uncertainty and this can be different on a case by case basis.

3.3 Boundedness of estimation error under high gain observers in the presence of uncertainty

We next present certain assumptions that enable state estimation and would be required for a stabilizing output feedback control design. The subsequent FDI design invokes these same assumptions to establish the ability of the proposed approach the achieve FDI. To this end, in this section we first utilize a high gain observer and establish boundedness of the estimation error in the presence of uncertainty. Consider the system of Eq. 3.1 under fault free conditions, satisfying Assumptions 3.1-3.3. Assumption 3.1 simply requires the dynamics system to be well behaved, and is satisfied by almost all systems of practical interest. Assumption 3.2, on the other hand, relies on an effective control system being in place for the system under consideration,

again a property naturally satisfied by all systems of practical interest (it is meaningful to design FDI mechanisms only for a system that is otherwise operating under a well designed control system i.e., in the absence of faults, the controller can meet the desired control objective including guaranteeing system stability).

Assumption 3.1. The functions $f : \mathbb{R}^n \to \mathbb{R}^n$, $g_i : \mathbb{R}^n \to \mathbb{R}^n$, $i = 1, \ldots, m$, $\theta(x, u, t) : \mathbb{R}^n \to \mathbb{R}^n$ and $h : \mathbb{R}^n \to \mathbb{R}^p$ are $\mathcal{C}^1$ functions on their domains of definition i.e. their derivatives exist and are continuous, and $f(0) = 0$.

Assumption 3.2. For the system of Eq. 3.1, there exists a positive definite $\mathcal{C}^2$ function $V : \mathbb{R}^n \to \mathbb{R}$ such that for any $x \in \Omega_c := \{x \in \mathbb{R}^n : V(x) \leq c\}$, where c is a positive real number, the following inequality holds:

$$L_f V(x) + L_g V(x) u(x) + L_\theta V(x) \leq -\alpha(V(x)) \tag{3.2}$$

where $L_g V(x) = [L_{g_1} V(x), \ldots, L_{g_m} V(x)]$, $u : \Omega_c \to \mathcal{U}$ is a state feedback control law and α is a class $\mathcal{K}$ function.

Remark 3.2. It is recognized that there exists no general procedure for construction of robust control Lyapunov functions (RCLFs) for nonlinear uncertain systems of the form of Eq. 3.1. However, for several classes of nonlinear uncertain systems, such a procedure exists (see e.g., Freeman and Kokotovic (2008) for further details). Also, if CLFs are used within an appropriately designed robust control law Mahmood *et al.* (2008a) can be verified to be RCLFs. Beside these results, it is more important to note that Assumption 3.2 simply states that a control design is in place to handle the uncertainly in the system, and does not require the knowledge of the specific Lyapunov function for the FDI design.

Assumption 3.3 below invokes the property that the system must be observable to enable output feedback stabilization.

Assumption 3.3. There exist integers ω_i, $i = 1, \ldots, p$, with $\sum_{i=1}^{p} \omega_i = n$, and a coordinate transformation $\zeta = T(x,u,t) = T'(x,u) + T_\theta(x,u,t)$ such that if $u = \bar{u}$, where $\bar{u} \in \mathcal{U}$ is a constant vector, then the representation of the system of Eq. 3.1 in the ζ coordinate takes the following form:

$$\begin{aligned} \dot{\zeta} &= A\zeta + B\phi(\zeta, \bar{u}) + \eta(\zeta, \bar{u}, t) \\ y &= C\zeta \end{aligned} \tag{3.3}$$

where $\zeta = [\zeta_1, \ldots, \zeta_p]^{\mathrm{T}} \in \mathbb{R}^n$, $A = \text{blockdiag}[A_1, \ldots, A_p]$, $B = \text{blockdiag}[B_1, \ldots, B_p]$, $C = \text{blockdiag}[C_1, \ldots, C_p]$, $T' = \phi = [\phi_1, \ldots, \phi_p]^{\mathrm{T}}$, $T_\theta = \eta = [\eta_1, \ldots, \eta_p]^{\mathrm{T}}$, $\zeta_i = [\zeta_{i,1}, \ldots, \zeta_{i,\omega_i}]^{\mathrm{T}}$, $A_i = \begin{bmatrix} 0 & I_{\omega_i - 1} \\ 0 & 0 \end{bmatrix}$, with $I_{\omega_i - 1}$ being a $(\omega_i - 1) \times (\omega_i - 1)$ identity matrix, $B_i = [0^{\mathrm{T}}_{\omega_i - 1},\ 1]^{\mathrm{T}}$, with $0_{\omega_i - 1}$ being a vector of zeros of dimension $\omega_i - 1$, $C_i = [1,\ 0^{\mathrm{T}}_{\omega_i - 1}]$, $\phi_i(\zeta, \bar{u}) = \phi_{i,\omega_i}(\zeta, \bar{u})$, with $\phi_{i,\omega_i}(\zeta, \bar{u})$ defined through the successive differentiation of $h_i(x)$: $\phi_{i,1}(\zeta, \bar{u}) = h_i(x)$ and $\phi_{i,j}(\zeta, \bar{u}) = \frac{\partial \phi_{i,j-1}}{\partial x}[f(x) + g(x)\bar{u}]$ and $\eta_i(\zeta, \bar{u}, t) = \eta_{i,\omega_i}(\zeta, \bar{u}, t)$, with $\eta_{i,\omega_i}(\zeta, \bar{u}, t)$ defined: $\eta_{i,1}(\zeta, \bar{u}, t) = 0$ and $\eta_{i,j}(\zeta, \bar{u}, t) = \frac{\partial \phi_{i,j-1}}{\partial x}[\theta(x,u,t)]$. Furthermore, $T' : \mathcal{X} \times \mathcal{U} \to \mathbb{R}^n$, $T_\theta : \mathcal{X} \times \mathcal{U} \to \mathbb{R}^n$, $T'^{-1} : \mathcal{X} \times \mathcal{U} \to \mathbb{R}^n$, and $T_\theta^{-1} : \mathcal{X} \times \mathcal{U} \to \mathbb{R}^n$ are $\mathcal{C}^1$ functions on their domains of definition, η denotes the uncertainty in the new coordinate system and $\|\eta(\zeta, u, t)\| \leq \bar{\eta}$, where $\bar{\eta}$ is a known positive constant.

We next describe the utilized high-gain observer formulation subject to sample-and-hold control. In particular, in the closed-loop system, the input is prescribed at discrete times $t_k = k\Delta$, $k = 0, \ldots, \infty$, with Δ being the hold-time of the control

action. The observer dynamics for $t \in [t_k, t_{k+1})$, are designed as follows (Du and Mhaskar (2014)):

$$\begin{aligned} \dot{\hat{\zeta}}(t) &= A\hat{\zeta}(t) + B\phi_0(\hat{\zeta}(t), u(t_k)) + H(y(t) - C\hat{\zeta}(t)) \\ \hat{\zeta}(t_k) &= T'(\hat{x}(t_k), u(t_k)) \end{aligned} \tag{3.4}$$

where $\hat{x}$ and $\hat{\zeta}$ denote the estimates of x and ζ, respectively, $H = \text{blockdiag}[H_1, \ldots, H_p]$ is the observer gain, $H_i = \left[\frac{a_{i,1}}{\varepsilon}, \ldots, \frac{a_{i,\omega_i}}{\varepsilon^{\omega_i}}\right]^{\mathrm{T}}$, with $s^{\omega_i} + a_{i,1}s^{\omega_i - 1} + \cdots + a_{i,\omega_i} = 0$ being a Hurwitz polynomial where where $i = 1, \ldots, p$ and ε being a positive constant to be specified, $\hat{x}(t_k) = T'^{-1}(\hat{\zeta}(t_k^-), u(t_{k-1})))$ for $k = 1, \ldots, \infty$, and ϕ_0 is the nominal model of ϕ. The initial state of the observer is denoted by $\hat{\zeta} := \hat{\zeta}(0)$. $\hat{\zeta}$ is re-initialized at the discrete times to account for the possible discrete changes in the input and ensuring that the resulting state estimates remain continuous.

Remark 3.3. In contrast to previous assumptions, the present assumption is specific to systems under control and hold implementation. This turns the control action into a fixed parameter. Thus the Assumption 3.3 provides a modified version of input-output normal form in the presence of uncertainty, and does not require the dynamics to be affine in the unmeasured states. This, along with existing techniques to handle measurement noise (see e.g., Du and Mhaskar (2014) and Ahrens and Khalil (2009)) significantly enhances the applicability of the designed observer, and in turn, the FDI mechanism. Benefiting from these relaxed assumptions on the dynamic system of Eq. 3.1 the proposed FDI design allow inclusion of a more general class of nonlinear systems compared to the previous designs (see e.g., Yan and Edwards (2007) and Zhang *et al.* (2010b)). In particular, no specific form for $f(x)$ and $g(x)$ (as in Zhang *et al.* (2010b)), nor any additional assumptions on $\theta(x, u, t)$ (as in Yan and Edwards

(2007)) are imposed.

Another requirement is the global boundedness of ϕ_0 formalized in Assumption 3.4 below (with a choice of zero readily satisfying the assumption).

Assumption 3.4. Findeisen *et al.* (2003) $\phi_0(\zeta, u)$ is a $\mathcal{C}^0$ function on its domain of definition and globally bounded in ζ.

Preparatory to the presentation of results on the convergence of the observer, we first state an important property of the scaled estimation error. To this end, let $D = \text{blockdiag}[D_1, \ldots, D_p]$, where $D_i = \text{diag}[\varepsilon^{\omega_i - 1}, \ldots, 1]$, and define the scaled estimation error $e = D^{-1}(\zeta - \hat{\zeta}) \in \mathbb{R}^n$. For $t \in [t_k, t_{k+1})$, the scaled estimation error evolves as follows:

$$\begin{aligned} \varepsilon \dot{e} &= A_0 e + \varepsilon B[\phi(\zeta, u(t_k)) - \phi_0(\hat{\zeta}, u(t_k))] + \varepsilon \eta(\zeta, u(t_k), t) \\ e(t_k) &= D^{-1}[T(x(t_k), u(t_k), t) - T'(\hat{x}(t_k), u(t_k), t)] \end{aligned} \tag{3.5}$$

where $A_0 = \text{blockdiag}[A_{0,1}, \ldots, A_{0,p}]$, $A_{0,i} = [a_i, b_i]$, $a_i = [-a_{i,1}, \ldots, -a_{i,\omega_i}]^{\mathrm{T}}$, and $b_i = [I_{\omega_i - 1}, 0_{\omega_i - 1}]^{\mathrm{T}}$.

Applying the change of time variable $\tau = \frac{t}{\varepsilon}$ and setting $\varepsilon = 0$, the boundary-layer system is given by

$$\frac{de}{d\tau} = A_0 e \tag{3.6}$$

Note that Eq. 3.6 corresponds to the boundary layer system, not the actual system. Thus $\varepsilon = 0$ simply defines the boundary layer system. However, for the observer that is utilized to estimate the system states, ε must be positive to have a finite gain and a feasible estimation scheme.

For the boundary-layer system, we define a Lyapunov function $W(e) = e^{\mathrm{T}} P_0 e$,

where P_0 is the symmetric positive definite solution of the Lyapunov function $A_0^{\mathrm{T}}P_0 + P_0A_0 = -I$. Let $\lambda_{\min}$ and $\lambda_{\max}$ denote the minimum and maximum eigenvalues of P_0, respectively. Proposition 3.1 below is similar to Proposition 1 in Du and Mhaskar (2014) and result obtained in Atassi and Khalil (1999) and hence stated without proof.

Proposition 3.1. Consider the system of Eq. 3.1, for which Assumptions 3.1, 3.3 and 3.4 hold. If $x_0 := x(0) \in \Omega_b$, where $0 < b < c$, then given $b' \in (b, c)$, there exists a finite time t_e, independent of ε, such that $x(t) \in \Omega_{b'}$ for all $t \in [0, t_e]$. Furthermore, there exists $\sigma > 0$, independent of ε, such that for any $e(t) \in \mathcal{W}_o := \{e \in \mathbb{R}^n : W(e) \geq \sigma\varepsilon^2\}$ and $x(t) \in \Omega_c$, $\dot{W} \leq -\frac{1}{2\varepsilon}\|e\|^2$.

Theorem 3.1 formalizes the convergence property of observer design and stability of closed loop system in the presence of uncertainty.

Theorem 3.1. *Consider the system of Eq. 3.1, for which Assumptions 3.1-3.4 hold, under a stabilizing control law u. Given any* $0 < b < c$, $d > 0$, $d' > 0$ *and* $\bar{\theta}$, *there exist* $\Delta^*(\bar{\theta}) > 0$ *and* $\varepsilon^*(\bar{\theta}) > 0$ *such that if* $\Delta \in (0, \Delta^*(\bar{\theta})]$, $\varepsilon \in (0, \varepsilon^*(\bar{\theta})]$, *and* $x_0 \in \Omega_b$, *then 1) there exists an integer* $k' > 0$ *such that* $\|\hat{x}(t_k) - x(t_k)\| \leq d'$ $\forall$ $t_k \geq t_{k'}$, *and 2)* $x(t) \in \Omega_c$ $\forall$ $t \geq 0$ *and* $\limsup_{t\to\infty} \|x(t)\| \leq d$.

Proof. The proof is divided into two parts. In the first part, we establish that the scaled estimation error converges to a bounded region in a finite time. To this end, we show that given $e_b(\bar{\theta}) > 0$, which is to be determined in the second part, there exists $\varepsilon^* > 0$ such that if $\varepsilon \in (0, \varepsilon^*(\bar{\theta})]$ and $\Delta \in (0, t_e]$, then the scaled estimation error $e(t_k^-)$ enters $\mathcal{E} := \{e \in \mathbb{R}^n : \|e\| \leq e_b(\bar{\theta})\}$ no later than the time t_e defined in Proposition 3.1, and stays in $\mathcal{E}$ thereafter as long as $x(t)$ remains in Ω_c. In the

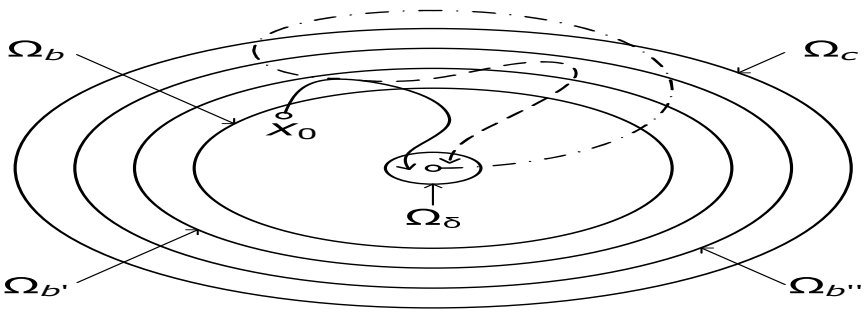


Figure 3.1: Schematic of the stability region, the evolution of the closed-loop state trajectories under fault-free (solid line) and faulty (dashed line) conditions, and the state estimate converging to its true value (dash-dotted line). The notation Ω_c denotes the stability region obtained under state feedback control (Du and Mhaskar (2014)).

second part of the proof, we show that boundedness of the scaled estimation error, e, guaranties boundedness of the estimation error, $\hat{x} - x$. To this end, we establish that for any $d > 0$ and $d' > 0$, there exists $e_b^* > 0$ and $\Delta^*(\bar{\theta}) > 0$ such that if $e(t_{k'}^-) \in \mathcal{E}$, $e_b \in (0, \min\{e_b^*(\bar{\theta}), \frac{d'-\bar{\theta}\Delta}{L_2}\}]$ where L_2 is a positive constant to be determined in the proof, and $\Delta \in (0, \Delta^*(\bar{\theta})]$, then $\|\hat{x}(t_k) - x(t_k)\| \leq d' \ \forall \ t_k \geq t_{k'}$, $x(t) \in \Omega_c \ \forall \ t \geq t_{k'}$, and $\limsup_{t\to\infty} \|x(t)\| \leq d$ (see the solid line in Figure 3.1)..

Consider $\Delta \in (0, \Delta_1(\bar{\theta})]$ and $\varepsilon \in (0, \varepsilon_1(\bar{\theta})]$, where $\Delta_1(\bar{\theta}) = t_e$ and $\varepsilon_1(\bar{\theta}) = \sqrt{\frac{\gamma}{\sigma}}$, with $0 < \gamma < \min_{\|e\|=e_b} W(e)$. In order to show that $e(t_k^-)$ converges to $\mathcal{E}$, we only need to show that it converges to $\mathcal{W}_i := \{e \in \mathbb{R}^n : W(e) \leq \sigma\varepsilon^2\}$.

Part 1: We first show that $e(t_k^-)$ reaches $\mathcal{W}_i$ no later than the time t_e. Let N be the largest integer such that $N\Delta \leq t_e$. It follows from Proposition 3.1 that if $t_{k+1} \leq t_e$, $k = 0, \ldots, N-1$, then for any $e \in \mathcal{W}_o$ and $t \in [t_k, t_{k+1})$, we have

$$\dot{W} \leq -\frac{1}{2\lambda_{\max}\varepsilon} W \tag{3.7}$$

It follows that

$$W(e(t_{k+1}^-)) \leq e^{-\frac{\Delta}{2\lambda_{\max}\varepsilon}} W(e(t_k)) \tag{3.8}$$

Let $\omega_{\max} = \max_{i=1,\ldots,p}\{\omega_i\}$. Since $T(x,u)$, $T'(x,u)$, $T^{-1}(\zeta,u)$ and $T'^{-1}(\zeta,u)$ are locally Lipschitz in x and ζ, respectively, and

$$\begin{aligned} e(t_k) &= D^{-1}[\zeta(t_k) - \hat{\zeta}(t_k)] = D^{-1}[T(x(t_k),u(t_k)) - T'(\hat{x}(t_k),u(t_k))] \\ &= D^{-1}[T'(x(t_k),u(t_k)) + T_\theta(x(t_k),u(t_k)) - T'(\hat{x}(t_k),u(t_k))] \end{aligned} \tag{3.9}$$

There exists $L_1, L_2 > 0$ such that the following equation holds:

$$\begin{aligned} \|e(t_k)\| &\leq L_1 L_2 \max\{1,\varepsilon^{1-\omega_{\max}}\} \times \max\{1,\varepsilon^{\omega_{\max}-1}\}\|e(t_k^-)\| + L_1 \max\{1,\varepsilon^{1-\omega_{\max}}\}\bar{T}_\theta \\ &+ L_1 \max\{1,\varepsilon^{1-\omega_{\max}}\}\bar{\theta}\Delta = L_1 L_2 \beta_1(\varepsilon)\|e(t_k^-)\| + L_1\beta_1(\varepsilon)\bar{T}_\theta + L_1\beta_1(\varepsilon)\bar{\theta}\Delta \end{aligned} \tag{3.10}$$

where $x(t_k) = T'^{-1}(\zeta(t_{k-1}),u(t_{k-1})) + T_{dev}^{-1}(\zeta(t_{k-1}),\, T_{dev}^{-1}(\zeta(t_{k-1}),u(t_{k-1})) = \int_{t_{k-1}}^{t_k}\theta(x,u,t)d\tau$ denotes the effect of the uncertainty on x, $\|T_\theta(x(t_k),u(t_k))\| \leq \bar{T}_\theta$, and $\beta_1(\varepsilon) = \varepsilon^{(\omega_{\max}-1)\mathrm{sgn}(\varepsilon-1)}$. Let $\tilde{L}_1 = L_1 L_2$. It follows from Eqs. 3.8 and 3.10 that if $e(t) \in \mathcal{W}_o$ for all $t \in [t_k, t_{k+1})$, then the following equation holds:

$$\begin{aligned} W(e(t_{k+1})) \leq{}& \frac{\lambda_{\max}}{\lambda_{\min}}\tilde{L}_1^2[\beta_1(\varepsilon)]^2 e^{-\frac{\Delta}{2\lambda_{\max}\varepsilon}} W(e(t_k)) + \lambda_{\max}L_1^2[\beta_1(\varepsilon)]^2\bar{T}_\theta^2 \\ &+ \lambda_{\max}L_1^2[\beta_1(\varepsilon)][\bar{\theta}\Delta]^2 + 2\lambda_{\max}\tilde{L}_1^2 L_1\beta_1^2(\varepsilon)\|e(t_k^-)\|\bar{T}_\theta \\ &+ 2\lambda_{\max}L_1^3\beta_1^2(\varepsilon)\|e(t_k^-)\|\ \bar{\theta}\Delta + 2\lambda_{\max}L_1^2\beta_1^2(\varepsilon)\bar{T}_\theta\bar{\theta}\Delta \end{aligned} \tag{3.11}$$

It follows form Eq. 3.7 that once $e(t)$ reaches $\mathcal{W}_i$, it stays there at least until the end of the same time interval. Since $T(x,u)$ is continuous, for any $x_0 \in \Omega_b$ and $\hat{x}_0 \in \mathcal{Q}$,

there exists $K_1 > 0$ such that

$$\|e(0)\| \leq K_1 \beta_2(\varepsilon) \tag{3.12}$$

where $\beta_2(\varepsilon) = \max\{1, \varepsilon^{1-\omega_{\max}}\}$ and as a result, using Eq. 3.10, we get

$$\begin{aligned}\|e(t_k)\| \leq\ & (\tilde{L}_1\beta_1(\varepsilon))^N (K_1\beta_2(\varepsilon))^N + (\tilde{L}_1\beta_1(\varepsilon))^{N-1}(L_1\beta_1(\varepsilon)\bar{T}_\theta + L_1\beta_1(\varepsilon)\bar{\theta}\Delta) \\ & + (\tilde{L}_1\beta_1(\varepsilon))^{N-2}(L_1\beta_1(\varepsilon)\bar{T}_\theta + L_1\beta_1(\varepsilon)\bar{\theta}\Delta) + \ldots \\ & + (L_1\beta_1(\varepsilon)\bar{T}_\theta + L_1\beta_1(\varepsilon)\bar{\theta}\Delta)\end{aligned} \tag{3.13}$$

To guarantee that $e(t_k^-)$ reaches $\mathcal{W}_i$ by the time t_N, it is required that the following equation hold:

$$\begin{aligned}&\frac{\lambda_{\max}}{\lambda_{\min}}\tilde{L}_1^2[\beta_1(\varepsilon)]^2 e^{-\frac{\Delta}{2\lambda_{\max}\varepsilon}} W(e(t_k)) + \lambda_{\max}L_1^2[\beta_1(\varepsilon)]^2\bar{T}_\theta^2 + \lambda_{\max}L_1^2[\beta_1(\varepsilon)]^2[\bar{\theta}\Delta]^2 \\ &+ 2\lambda_{\max}\tilde{L}_1^2 L_1\beta_1^2(\varepsilon)\|e(t_k^-)\|\bar{T}_\theta + 2\lambda_{\max}L_1^3\beta_1^2(\varepsilon)\|e(t_k^-)\|\ \bar{\theta}\Delta \\ &+ 2\lambda_{\max}L_1^2\beta_1^2(\varepsilon)\bar{T}_\theta\bar{\theta}\Delta \leq \sigma\varepsilon^2\end{aligned} \tag{3.14}$$

By using Eq. 3.13, the following inequality holds for the left hand side of Eq. 3.14:

$$\begin{aligned}&\frac{\lambda_{\max}}{\lambda_{\min}}\tilde{L}_1^2[\beta_1(\varepsilon)]^2 e^{-\frac{\Delta}{2\lambda_{\max}\varepsilon}} W(e(t_k)) + \lambda_{\max}L_1^2[\beta_1(\varepsilon)]^2\bar{T}_\theta^2 + \lambda_{\max}L_1^2[\beta_1(\varepsilon)]^2[\bar{\theta}\Delta]^2 \\ &+ 2\lambda_{\max}\tilde{L}_1^2 L_1\beta_1^2(\varepsilon)\|e(t_k^-)\|\bar{T}_\theta + 2\lambda_{\max}L_1^3\beta_1^2(\varepsilon)\|e(t_k^-)\|\ \bar{\theta}\Delta \\ &+ 2\lambda_{\max}L_1^2\beta_1^2(\varepsilon)\bar{T}_\theta\bar{\theta}\Delta \leq [\frac{\lambda_{\max}}{\lambda_{\min}}\tilde{L}_1^2[\beta_1(\varepsilon)]^2 e^{-\frac{\Delta}{2\lambda_{\max}\varepsilon}}]^N \lambda_{max}K_1^2[\beta_2(\varepsilon)]^2 \\ &+ \Gamma([\beta_1(\varepsilon)]^N[\beta_2(\varepsilon)]^N, [\beta_1(\varepsilon)]^s\bar{T}_\theta^q, [\beta_1(\varepsilon)]^n[\bar{\theta}\Delta]^z)^z)\end{aligned} \tag{3.15}$$

where Γ is continuous in ε and $s \geq q$ and $n \geq z$. Therefore if the following inequality

holds, then Eq. 3.14 holds too:

$$\begin{aligned}&\frac{1}{\varepsilon^2}\frac{\lambda_{\max}}{\lambda_{\min}}\tilde{L}_1^2[\beta_1(\varepsilon)]^2 e^{-\frac{\Delta}{2\lambda_{\max}\varepsilon}}]^N \lambda_{max} K_1^2[\beta_2(\varepsilon)]^2 \\ &+\frac{1}{\varepsilon^2}\Gamma([\beta_1(\varepsilon)]^N[\beta_2(\varepsilon)]^N, [\beta_1(\varepsilon)]^s\bar{T}_\theta^q, [\beta_1(\varepsilon)]^n[\bar{\theta}\Delta]^z) \leq \sigma\end{aligned} \tag{3.16}$$

Since the left-hand side of the above inequality is continuous in ε and tends to zero as ε tends to 0, there exists $\varepsilon_2(\bar{\theta}) > 0$ such that if $\varepsilon \in (0, \varepsilon_2(\bar{\theta})]$, then Eq. 3.14 holds.

We then show that after the scaled estimate error $e(t_k^-)$ reaches $\mathcal{W}_i$, it stays there as long as $x(t)$ stays in Ω_c. Note that given $e(t_k^-) \in \mathcal{W}_i$, it is possible that $e(t_k)$ goes outside $\mathcal{W}_i$ due to the re-initialization to the system state and its estimate in the ζ coordinate. It follows from Eq. 3.10 that if $e(t_k^-) \in \mathcal{W}_i$, then $\|e(t_k)\| \leq \tilde{L}_1\beta_1(\varepsilon)e_b + L_1\beta_1(\varepsilon)\bar{T}_\theta + L_1\beta_1(\varepsilon)\bar{\theta}\Delta$.

To guarantee that $e(t_{k+1}^-)$ stays in $\mathcal{W}_i$, it is required that the following equation hold:

$$e^{-\frac{\Delta}{2\lambda_{\max}\varepsilon}}\lambda_{\max}\tilde{L}_1^2[\beta_1(\varepsilon)]^2 e_b^2 + L_1\beta_1(\varepsilon)\bar{T}_\theta + L_1\beta_1(\varepsilon)\bar{\theta}\Delta \leq \sigma\varepsilon^2 \tag{3.17}$$

It can be shown that there exists $\varepsilon_3(\bar{\theta}) > 0$ such that if $\varepsilon \in (0, \varepsilon_3(\bar{\theta})]$, then Eq. 3.17 holds.

Part 2: We first show that if the system state resides within a subset of Ω_c and the scaled estimation error is sufficiently small, then the state estimate also resides within Ω_c. It follows from the first part of the proof that we have

$$\begin{aligned}\|x - \hat{x}\| &= \|T^{-1}(\zeta, u) - T'^{-1}(\hat{\zeta}, u)\| = \|T'^{-1}(\zeta, u) + T_{dev}^{-1}(\zeta, u) - T'^{-1}(\hat{\zeta}, u)\| \\ &\leq L_2\beta_3(\varepsilon)\|e\| + \bar{\theta}\Delta \leq L_2\beta_3(\varepsilon_1)\|e\| + \bar{\theta}\Delta\end{aligned} \tag{3.18}$$

where $\beta_3(\varepsilon) = \max\{1, \varepsilon^{\omega_{\max}-1}\}$. It can be shown that given $0 < \delta_1 < \delta_2$, there exists

$\tilde{e} > 0$ such that if $e_b \in (0, \tilde{e}]$, then $V(x) \leq \delta_1$ implies $V(\hat{x}) \leq \delta_2$. It follows from Proposition 3.1 that given $b' \in (b, c)$, we have that $x(t_{k'}) \in \Omega_{b'}$. Therefore, there exists $e_{b,1} > 0$ such that if $e_b \in (0, e_{b,1}(\bar{\theta})]$, then $\hat{x}(t_{k'}) \in \Omega_c$.

We then show the existence of $e_b^*(\bar{\theta}) > 0$ and $\Delta^*(\bar{\theta}) > 0$ such that if $e_b \in (0, e_b^*(\bar{\theta})]$ and $\Delta \in (0, \Delta^*(\bar{\theta})]$, then any state trajectory originating in $\Omega_{b'}$ at time $t_{k'}$ converges to a closed ball of radius d around the origin. Since $V(x)$ is a continuous function of the state, one can find a positive real number $\delta < b'$ such that $V(x) \leq \delta$ implies $\|x\| \leq d$. Let $\hat{\delta}$ be a positive real number such that $0 < \hat{\delta} < \delta$. If $e_b \in (0, e_{b,1}(\bar{\theta})]$, the state estimate at time $t_{k'}$ can either be such that $\hat{\delta} < V(\hat{x}(t_{k'})) \leq c$ or $V(\hat{x}(t_{k'})) \leq \hat{\delta}$.

Case 1: Consider $\hat{x}(t_k) \in \Omega_c \backslash \Omega_{\hat{\delta}}$. Let $\dot{V}(x, u) = L_f V(x) + L_g V(x) u + L_\theta V(x)$. For this case, we have $\dot{V}(\hat{x}(t_k),\, u(t_k)) \leq -\alpha(V(\hat{x}(t_k))) < -\alpha(\hat{\delta})$. It follows from the continuity properties of $f(\cdot)$, $g(\cdot)$, $\theta(\cdot)$ and $V(\cdot)$ that $L_f V(\cdot)$, $L_g V(\cdot)$ and $L_\theta V(\cdot)$ are locally Lipschitz on the domain of interest. Therefore, there exists $L_3 > 0$ such that

$$|\dot{V}(x(t_k), u(t_k)) - \dot{V}(\hat{x}(t_k), u(t_k))| \leq L_3 \|x(t_k) - \hat{x}(t_k)\| \quad \leq L_2 L_3 \beta_3(\varepsilon_1) \|e(t_k^-)\| + L_3 \bar{\theta} \Delta \tag{3.19}$$

Since the functions $f(\cdot)$, $g(\cdot)$ and $\theta(\cdot)$ are continuous, u is bounded, and $\Omega_{b'}$ is bounded, one can find $K_2 > 0$ such that $\|x(t) - x(t_k)\| \leq K_2 \Delta$ for any $\Delta \in (0, \Delta_1]$, $x(t_k) \in \Omega_{b'}$ and $t \in [t_k, t_k + \Delta)$. It follows that $\forall\ t \in [t_k, t_k + \Delta)$, the following equation holds:

$$\dot{V}(x(t)) < -\alpha(\hat{\delta}) + L_3 K_2 \Delta + L_2 L_3 \beta_3(\varepsilon_1(\bar{\theta})) \|e(t_k^-)\| + L_3 \bar{\theta} \Delta \tag{3.20}$$

Consider $\Delta \in (0, \Delta_2(\bar{\theta})]$, where $\Delta_2 = \frac{\alpha(\hat{\delta})}{3 L_3 K_2}$, and $e_b \in (0, e_{b,2}(\bar{\theta})]$, where $e_{b,2}(\bar{\theta}) =$

$\frac{\alpha(\hat{\delta})-L_3\bar{\theta}\Delta}{3L_2L_3\beta_3(\varepsilon_1(\bar{\theta}))}$. Then, we have

$$\dot{V}(x(t)) < -\frac{1}{3}\alpha(\hat{\delta}) < 0 \tag{3.21}$$

Since $\dot{V}(x(t))$ remains negative over $[t_k, t_k + \Delta)$, $x(t)$ remains in Ω_c over the same time interval, and $V(x(t_k + \Delta)) < V(x(t_k))$.

If $\hat{x}(t_{k'}) \in \Omega_c \backslash \Omega_{\hat{\delta}}$, we have $\dot{V}(x(t)) < 0$ over $[t_{k'}, t_{k'} + \Delta)$. It follows that $\hat{x}(t_{k'+1}) \in \Omega_c$ for $e_b \in (0, e_{b,1}(\bar{\theta})]$. Similarly, it can be shown that for $t_k > t_{k'}$, $\dot{V}(x(t))$ remains negative until $\hat{x}(t_k)$ reaches $\Omega_{\hat{\delta}}$.

Case 2: Consider $\hat{x}(t_k) \in \Omega_{\hat{\delta}}$. For this case, it is established in the proof of the second part of Theorem 1 in Du and Mhaskar (2014), that there exist $e_{b,3}(\bar{\theta}) > 0$ and $\Delta_3(\bar{\theta}) > 0$ such that if $e_b \in (0, e_{b,3}(\bar{\theta})]$ and $\Delta \in (0, \Delta_3(\bar{\theta})]$, we have $x(t_{k+1}) \in \Omega_\delta$ and as a result $\hat{x}(t_{k+1}) \in \Omega_c$ for $e_b \in (0, e_{b,1}(\bar{\theta})]$).

For $e_b \in (0, e_b^*(\bar{\theta})]$ and $\Delta \in (0, \Delta^*(\bar{\theta})]$, where $e_b^* = \min\{e_{b,1}(\bar{\theta}), e_{b,2}(\bar{\theta}), e_{b,3}(\bar{\theta})\}$ and $\Delta^*(\bar{\theta}) = \min\{\Delta_1(\bar{\theta}), \Delta_2(\bar{\theta}), \Delta_3(\bar{\theta})\}$, it can be shown by iteration that any state trajectory originating in $\Omega_{b'}$ at time $t_{k'}$ converges to the set Ω_δ, and hence converges to the closed ball of radius d around the origin. Furthermore, if $e_b \leq \frac{d'-\bar{\theta}\Delta}{L_2}$, it follows from Eq. 3.18 that $\|\hat{x}(t_k) - x(t_k)\| \leq d' \ \forall \ t_k \geq t_{k'}$.

In summary, it is established that given any $0 < b < c$, $d > 0$ and $d' > 0$, there exist $\Delta^*(\bar{\theta}) > 0$ and $\varepsilon^*(\bar{\theta}) > 0$ such that if $\Delta \in (0, \Delta^*(\bar{\theta})]$, $\varepsilon \in (0, \varepsilon^*(\bar{\theta})]$, and $x_0 \in \Omega_b$, then 1) $\|\hat{x}(t_k) - x(t_k)\| \leq d' \ \forall \ t_k \geq t_{k'}$, and 2) $x(t) \in \Omega_c \ \forall \ t \geq 0$ and $\limsup_{t\to\infty} \|x(t)\| \leq d$. This concludes the proof of Theorem 3.1. □

3.4 Fault detection and isolation mechanism

This section presents the proposed fault detection and isolation mechanism that utilizes the error convergence properties of the observers established in the previous section. We restrict our attention to scenarios where at most two, actuators and/or sensors, experiences a fault. By a direct application of the principles of combinatorics, this leads to $n_f = m + \frac{m(m-1)}{2} + p + \frac{p(p-1)}{2} + mp$ unique scenarios. Residuals (with the associated high gain observers) are next designed for each fault scenario in the same fashion as Du *et al.* (2013).

Before proceeding to present the FDI mechanism, we need to employ the following assumption:

Assumption 3.5. The systems state vector x remains bounded before and after fault occurrence i.e., there exist a positive constant d_g such that $\|x\| \leq d_g$, $\forall t > 0$.

Remark 3.4. Note that Assumption 3.5 states that the proposed FDI methodology remains applicable under any stabilizing output feedback controller implemented in a discrete fashion as long as the system state evolves within the compact set $\mathcal{X}$ before and after fault occurrence.

Specifically, the residual for a fault scenario is defined as the norm of the difference between the state prediction and the state estimate for the subsystem model that does not require the value of the corresponding actuator/sensor in the calculations (see Du *et al.* (2013) for more details). A residual dedicated to a particular fault scenario is designed so it is only insensitive to that particular fault scenario but sensitive to the other fault scenarios (we thus design so-called generalized residuals). To this end, let $\Theta_{f,i}$ denote the fault vector (sensor/and or actuator) for the ith fault scenario,

and $\bar{\Theta}_{f,i}$ the remaining fault variable vector (the remaining u_f and y_f variables). Specifically, let $u_{f,i}$ and $y_{f,i}$ denote the vectors of input and output variables subject to faults $\Theta_{f,i}$, respectively. Let $\bar{u}_{f,i}$ and $\bar{y}_{f,i}$ denote the vectors of the rest of the input and output variables, respectively. We next state an assumption that is required for the ability to detect and isolate sensor and actuator faults. In particular, the assumption requires that it be possible to 'observe' a particular variable (sensor or actuator) in more ways than one, to in turn be able to detect and isolate faults. This assumption is therefore necessary to achieve FDI.

Assumption 3.6. Du *et al.* (2013) Assumptions 3.3 and 3.4 hold for the system of Eq. 3.1, with $\bar{u}_{f,i}$ and $\bar{y}_{f,i}$ being the vectors of input and output variables, respectively, $i = 1, \ldots, n_f$ where n_f is the number of possible fault scenarios.

Under Assumption 3.6 and by choosing $\phi_0 = 0$, the state observer for the ith fault scenario is designed as follow :

$$\begin{aligned} \dot{\hat{\zeta}}^j &= A^j \hat{\zeta}^j + H^j(\bar{y}_{f,i} - C^j \hat{\zeta}^j) \\ \dot{\hat{\zeta}}^j(t_k) &= T'^j(\hat{x}^j(t_k), \bar{u}_{f,i}(t_k)) \end{aligned} \tag{3.22}$$

where j represents the jth observer with $j = 1, \ldots, p + \frac{p(p-1)}{2}$. To define residuals, we need to compute expected trajectories. To this end, we consider a subsystem of Eq. 3.1 for which the state variables are all of those such that no inputs in u_{f_i} appear on the right-hand side of the corresponding ODE's. Let x_{sub} denote the vector of state variables for the subsystem, and $\bar{x}_{sub,}$ the vector of the rest of the state variables.

Without loss of generality, the model of the subsystem can be described as follows:

$$\begin{aligned}\dot{x}_{sub,i} = f_{sub,i}([x_{sub,i}^{\mathrm{T}}, \bar{x}_{sub,i}^{\mathrm{T}}]^{\mathrm{T}}) + G_{sub,i}([x_{sub,i}^{\mathrm{T}}, \bar{x}_{sub,i}^{\mathrm{T}}]^{\mathrm{T}})\bar{u}_{f,i} \\ + \theta_{sub,i}([x_{sub,i}^{\mathrm{T}}, \bar{x}_{sub,i}^{\mathrm{T}}]^{\mathrm{T}}, \bar{u}_{f,i}, t)\end{aligned} \tag{3.23}$$

where $f_{sub,i}(\cdot)$, $G_{sub,i}(\cdot)$ and $\theta_{sub,i}(\cdot)$ are appropriately defined. For each faulty scenario, the expected system trajectory is computed using the known part of the system model and the state estimates generated by the jth observer that does not require values of the variables included in the fault vector $\Theta_{f,i}$. Specifically, for $t \in [t_{k-T}, t_k)$, a prediction model is designed as follows:

$$\dot{\tilde{x}}_{sub,i,j} = f_{sub,i}([\tilde{x}_{sub,i,j}^{\mathrm{T}}, \hat{\tilde{x}}_{sub,i,j}^{\mathrm{T}}]^{\mathrm{T}}) + G_{sub,i}([\tilde{x}_{sub,i,j}^{\mathrm{T}}, \hat{\tilde{x}}_{sub,i,j}^{\mathrm{T}}]^{\mathrm{T}})\bar{u}_{f,i} \tag{3.24}$$

where $\tilde{x}_{sub,i,j}$ is the state of the prediction model, $\hat{\tilde{x}}_{sub,i,j}$ is the estimate of $\bar{x}_{sub,i}$ provided by the jth observer, and T is the prediction horizon: $T = 1$ if $0 < t_k \leq t_{k'}$; $T = k - k'$ if $t_{k'} < t_k \leq t_{k'+T_p}$; and $T = T_p$ if $t_k > t_{k'+T_p}$, with a positive integer T_p being the prediction horizon after the initialization period. The initial condition for the prediction model is the state estimate at time t_{k-T}: $\tilde{x}_{sub,i,j}(t_{k-T}) = \hat{x}_{sub,i,j}(t_{k-T})$. Let $\tilde{x}_{sub,i,j}(t_k)$ denote the prediction for the state vector $x_{sub,i}$ at time t_k. By solving Eq. 3.24, the state prediction at time t_k is obtained. The residual for a particular actuator are defined as the norm of the difference between the state prediction and the state estimate for the subsystem that is not subject to that particular actuator. For the ith faulty scenario, the residual (at the time instance t_{k+1}) is defined as follows:

$$r_{i,k+1} = \|\tilde{x}_{sub,i,j}(t_{k+1}) - \hat{x}_{sub,i,j}(t_{k+1})\| \tag{3.25}$$

Remark 3.5. Note that for defining the corresponding residuals insensitive to actuator faults, the prediction model utilizes the state measurements, if they are available. If not, they must be replaced by state estimates computed by the observer that does not require knowledge of the prescribed input. For defining the corresponding residuals insensitive to simultaneous actuator and a particular sensor faults, the specific sensor measurements can no longer be used in the prediction model and they are replaced by state estimates generated by an observer that does not use the prescribed input nor the specific sensor measurement. This is the key feature that enables us to distinguish between actuator faults and simultaneous actuator and sensor faults.

Now we specify how the thresholds for the residuals need to be selected to enable FDI. Before proceeding to define thresholds corresponding to each residual, we need the following result, that determines the bounds on the residuals in the presence of uncertainty:

Lemma 3.1. Consider the system of Eq. 3.1, for which Assumptions 3.1-3.6 hold and residuals for each sample time are defined as in Eq. 3.25. Then under fault free condition, for $t_k \geq t_{k'}$ (see Theorem 3.1 for definition of k'):

$$r_{i,k+1} \leq L_{2,sub,i,j}\beta_3^i(\varepsilon_j)E_{s,sub,i,j} + \bar{\theta}_{sub,i}\Delta + E_p^i \tag{3.26}$$

where $E_{s,sub,i,j} = e^{-\alpha_{sub,i,j}\Delta}e_b^{*,i} + K_{B,sub,i,j}^i K_{\phi,,sub,i,j}^i \frac{1-e^{-\alpha_{sub,i,j}\Delta}}{\alpha_{sub,i,j}} + \frac{1-e^{-\alpha_{sub,i,j}\Delta}}{\alpha_{sub,i,j}}\bar{\eta}_{sub,i,j}$, sub and j refer to the corresponding subsystem and observer used for defining r_i, respectively, where i refers to ith residual, $e_b^{*,i}$ is the upper bound on $\|e_{sub,i,j}^i(t_k)\|$, K_{B_i} is spectrum bound of the matrix $\|B^i\|$, α_i is the spectrum bound of the matrix $\frac{A_0^i}{\varepsilon}$ and $E_p^i = M^i\Delta + K_{\tilde{e}}$, $\|(\tilde{f}_{sub,i} - f_{sub,i}([x_{sub,i}^{\mathrm{T}}, \bar{x}_{sub,i}^{\mathrm{T}}]^{\mathrm{T}})) + (\tilde{G}_{sub,i} - G_{sub,i}([x_{sub,i}^{\mathrm{T}}, \bar{x}_{sub,i}^{\mathrm{T}}]^{\mathrm{T}}))\bar{u}_{f,i}\| +$

$\bar{\theta}_{sub,i} \leq M^i$,

where $\tilde{f}_{sub,i} = f_{sub,i}([\tilde{x}^{\mathrm{T}}_{sub,i,j}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,i,j}]^{\mathrm{T}})$, $\tilde{G}_{sub,i} = G_{sub,i}([\tilde{x}^{\mathrm{T}}_{sub,i,j}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,i,j}]^{\mathrm{T}})$, $K_{\tilde{e}}$ is the upper bound on $\|\tilde{e}_{i,k}(t_k)\|$ which is constant, where $\tilde{e}_{i,k} = \tilde{x}_{sub,i,j} - x_{sub,i}$ and $\tilde{x}_{sub,i,j}$ is state of prediction model defined in Du *et al.* (2013).

Proof. By using triangular inequality, Eq. 3.25 turns to the following form:

$$\begin{aligned} r_{i,k+1} &= \|\tilde{x}_{sub,i,j}(t_{k+1}) - \hat{x}_{sub,i,j}(t_{k+1})\| \\ &\leq \|\tilde{x}_{sub,i,j}(t_{k+1}) - x_{sub,i}(t_{k+1})\| + \|x_{sub,i}(t_{k+1}) - \hat{x}_{sub,i,j}(t_{k+1})\| \\ &\leq \sup \|\tilde{x}_{sub,i,j}(t_{k+1}) - x_{sub,i}(t_{k+1})\| + \sup \|x_{sub,i}(t_{k+1}) - \hat{x}_{sub,i,j}(t_{k+1})\| \end{aligned} \tag{3.27}$$

Consider the prediction model corresponding to fault Θ_{f_i}:

$$\dot{\tilde{x}}_{sub,i,j} = f_{sub,i}([\tilde{x}^{\mathrm{T}}_{sub,i,j}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,i,j}]^{\mathrm{T}}) + G_{sub,i}([\tilde{x}^{\mathrm{T}}_{sub,i,j}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,i,j}]^{\mathrm{T}})\bar{u}_{f,i} \tag{3.28}$$

where $\tilde{x}_{sub,i,j}$ is the state of the prediction model, $\hat{\tilde{x}}_{sub,i,j}$ is the estimate of $\bar{x}_{sub,i}$ provided by the corresponding observer.

Defining $\tilde{e}_{i,k} = \tilde{x}_{sub,i,j} - x_{sub,i}$, we obtain:

$$\begin{aligned} \dot{\tilde{e}}_{i,k} = {} & f_{sub,i}([\tilde{x}^{\mathrm{T}}_{sub,i,j}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,i,j}]^{\mathrm{T}}) - f_{sub,i}([x^{\mathrm{T}}_{sub,i}, \bar{x}^{\mathrm{T}}_{sub,i}]^{\mathrm{T}}) + \\ & (G_{sub,i}([\tilde{x}^{\mathrm{T}}_{sub,i,j}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,i,j}]^{\mathrm{T}}) - G_{sub,i}([x^{\mathrm{T}}_{sub,i}, \bar{x}^{\mathrm{T}}_{sub,i}]^{\mathrm{T}}))\bar{u}_{f,i} \\ & - \theta_{sub,i}([x^{\mathrm{T}}_{sub,i}, \bar{x}^{\mathrm{T}}_{sub,i}]^{\mathrm{T}}, \bar{u}_{f,i}, t) \end{aligned} \tag{3.29}$$

For sake of brevity, we define $\tilde{f}_{sub,i} = f_{sub,i}([\tilde{x}^{\mathrm{T}}_{sub,i,j}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,i,j}]^{\mathrm{T}})$, $\tilde{G}_{sub,i} =$

$G_{sub,i}([\tilde{x}^{\mathrm{T}}_{sub,i,j}, \hat{\bar{x}}^{\mathrm{T}}_{sub,i,j}]^{\mathrm{T}})$. By integration, we get:

$$\begin{aligned}\tilde{e}_{i,k}(t_{k+1}) = &\int_{t_k}^{t_{k+1}} (\tilde{f}_{sub,i} - f_{sub,i}([x^{\mathrm{T}}_{sub,i}, \bar{x}^{\mathrm{T}}_{sub,i}]^{\mathrm{T}}))d\tau \\ &+ \int_{t_k}^{t_{k+1}} (\tilde{G}_{sub,i} - G_{sub,i}([x^{\mathrm{T}}_{sub,i}, \bar{x}^{\mathrm{T}}_{sub,i}]^{\mathrm{T}}))\bar{u}_{f,i} - \theta_{sub,i})d\tau + \tilde{e}_{i,k}(t_k)\end{aligned} \tag{3.30}$$

Under assumptions 3.3 and 3.4 and by applying the triangular inequality, we obtain:

$$\begin{aligned}\|\tilde{e}_{i,k}(t_{k+1})\| &\leq \int_{t_k}^{t_{k+1}} \|(\tilde{f}_{sub,i} - f_{sub,i}([x^{\mathrm{T}}_{sub,i}, \bar{x}^{\mathrm{T}}_{sub,i}]^{\mathrm{T}}))\|d\tau \\ &+ \int_{t_k}^{t_{k+1}} (\|(\tilde{G}_{sub,i} - G_{sub,i}([x^{\mathrm{T}}_{sub,i}, \bar{x}^{\mathrm{T}}_{sub,i}]^{\mathrm{T}}))\bar{u}_{f,i}\| + \|\theta_{sub,i}\|)d\tau + \|\tilde{e}_{i,k}(t_k)\| \\ &\leq M^i\Delta + K_{\tilde{e}}\end{aligned} \tag{3.31}$$

where $\|(\tilde{f}_{sub,i} - f_{sub,i}([x^{\mathrm{T}}_{sub,i}, \bar{x}^{\mathrm{T}}_{sub,i}]^{\mathrm{T}}))\| + \|(\tilde{G}_{sub,i} - G_{sub,i}([x^{\mathrm{T}}_{sub,i}, \bar{x}^{\mathrm{T}}_{sub,i}]^{\mathrm{T}}))\bar{u}_{f,i}\| + \bar{\theta}_{sub,i} \leq M^i$ and $\|\tilde{e}_{i,k}(t_k)\| \leq K_{\tilde{e}}$.

Now we need to determine the supremum for $\|x_{sub,i}(t_{k+1}) - \hat{x}_{sub,i,j}(t_{k+1})\|$. For $t \in [t_k, t_{k+1})$, the scaled estimation error corresponding to ith residual evolves as follows:

$$\dot{e^i} = \frac{1}{\varepsilon^i} A^i_0 e^i + B^i(\phi^i(\bar{u}_{f_i}) - \phi^i_0(\bar{u}_{f_i})) + \eta_i \tag{3.32}$$

The solution to the above equation gives

$$\begin{aligned}e^i(t_{k+1}) = &\int_{t_k}^{t_{k+1}} e^{\frac{A^i_0}{\varepsilon}(t_{k+1}-\tau)}[B^i(\phi^i(\zeta, \bar{u}_{f_i}, t) - \phi^i_0(\hat{\zeta}, \bar{u}_{f_i}, t)) + \eta_i(\zeta, u, t)]d\tau \\ &+ e^{-\frac{A^i_0}{\varepsilon}(t_k - t_{k+1})} e^i(t_k)\end{aligned} \tag{3.33}$$

Under Assumptions 3.3 and 3.4 and by applying the triangular inequality, we obtain:

$$\begin{aligned}
\|e^i(t_{k+1})\| &\leq \| \int_{t_k}^{t_{k+1}} e^{\frac{A_0^i}{\varepsilon}(t_{k+1}-\tau)}[B^i(\phi^i(\zeta,\bar{u}_{f_i},t)-\phi_0^i(\hat{\zeta},\bar{u}_{f_i},t))+\eta_i(\zeta,u,t)]d\tau\| \\
&+ \|e^{-\frac{A_0^i}{\varepsilon}(t_k-t_{k+1})}e^i(t_k)\| \\
&\leq \| \int_{t_k}^{t_{k+1}} e^{\frac{A_0^i}{\varepsilon}(t_{k+1}-\tau)}[B^i(\phi^i(\zeta,\bar{u}_{f_i},t)-\phi_0^i(\hat{\zeta},\bar{u}_{f_i},t))]d\tau\| \\
&+ \| \int_{t_k}^{t_{k+1}} e^{\frac{A_0^i}{\varepsilon}(t_{k+1}-\tau)}\eta_i(\zeta,u,t)]d\tau\| + \|e^{-\frac{A_0^i}{\varepsilon}(t_k-t_{k+1})}e^i(t_k)\| \\
&\leq \int_{t_k}^{t_{k+1}} e^{\frac{A_0^i}{\varepsilon}(t_{k+1}-\tau)}\|B^i\|\|(\phi^i(\zeta,\bar{u}_{f_i},t)-\phi_0^i(\hat{\zeta},\bar{u}_{f_i},t))\|d\tau \\
&+ \int_{t_k}^{t_{k+1}} e^{\frac{A_0^i}{\varepsilon}(t_{k+1}-\tau)}\|\eta_i(\zeta,u,t)\|d\tau + e^{-\frac{A_0^i}{\varepsilon}(t_k-t_{k+1})}\|e^i(t_k)\| \\
&\leq e^{-\alpha\Delta}e_b^{*,i} + K_{B_i}K_{\phi_i}\frac{1-e^{-\alpha_i\Delta}}{\alpha_i} + \frac{1-e^{-\alpha_i\Delta}}{\alpha_i}\bar{\eta}_i
\end{aligned} \tag{3.34}$$

where α_i, K_{B_i}, K_ϕ are the spectrum bound of the matrices $\frac{A_0^i}{\varepsilon}$, B^i, $(\phi^i(\zeta,\bar{u}_{f_i},t)-\phi_0^i(\hat{\zeta},\bar{u}_{f_i},t))$ respectively, and $e_b^{*,i}$ is the upper bound on $\|e^i(t_k)\|$. From 3.18, we have

$$\|x-\hat{x}\| = \|T_j^{-1}(\zeta,u) - T_j'^{-1}(\hat{\zeta},u)\| \leq L_{2,j}\beta_3(\varepsilon_j)E_{s,i} + \bar{\theta}_i\Delta \tag{3.35}$$

where $\beta_3(\varepsilon_j) = \max\{1, \varepsilon_j^{\omega_{\max}-1}\}$, $L_{2,j} > 0$, $E_s = e^{-\alpha\Delta}e_b^{*,i} + K_{B_i}K_{\phi_i}\frac{1-e^{-\alpha_i\Delta}}{\alpha_i} + \frac{1-e^{-\alpha_i\Delta}}{\alpha_i}\bar{\eta}_i$ and as a result,

$$\|x_{sub,i} - \hat{x}_{sub,i,j}\| \leq L_{2,sub,i,j}\beta_3(\varepsilon_j)E_{s,sub,i,j} + \bar{\theta}_{sub,i}\Delta \tag{3.36}$$

where $E_{s,sub,i,j} = e^{-\alpha_{sub,i,j}\Delta}\|e^i_{sub,i,j}(t_k)\| + K_{B,sub,i,j}K_{\phi,,sub,i,j}\frac{1-e^{-\alpha_{sub,i,j}\Delta}}{\alpha_{sub,i,j}} + \frac{1-e^{-\alpha_{sub,i,j}\Delta}}{\alpha_{sub,i,j}}\bar{\eta}_{sub,i,j}$.

Using Eq. 3.31 and 3.35, $r_{i,k+1}$ is bounded as below:

$$\begin{aligned} r_{i,k+1} &= \|\tilde{x}_{sub,i,j}(t_{k+1}) - \hat{x}_{sub,i,j}(t_{k+1})\| \\ &\leq \|\tilde{x}_{sub,i,j}(t_{k+1}) - x_{sub,i}(t_{k+1})\| + \|x_{sub,i}(t_{k+1}) - \hat{x}_{sub,i,j}(t_{k+1})\| \\ &\leq \sup \|\tilde{x}_{sub,i,j}(t_{k+1}) - x_{sub,i}(t_{k+1})\| + \sup \|x_{sub,i}(t_{k+1}) - \hat{x}_{sub,i,j}(t_{k+1})\| \\ &= L_{2,sub,i,j}\beta_3^i(\varepsilon_j)E_{s,sub,i,j} + \bar{\theta}_{sub,l}\Delta + E_p^i \end{aligned} \tag{3.37}$$

where $E_p = M^i\Delta + K_{\tilde{e}}$. This concludes proof of lemma 3.1. □

Having determined the bound, the threshold corresponding to each residual are picked as below:

$$\delta_i = L_{2,sub,i,j}\beta_3^i(\varepsilon_j)E_{s,sub,i,j} + \bar{\theta}_{sub,l}\Delta + E_p^i \tag{3.38}$$

It follows from Lemma 1 then that under fault free condition $r_{i,k+1} \leq \delta_i$.

Remark 3.6. Note that the threshold defined by Eq. 3.38 depends on the observer gain, H, since α_i represents the spectrum bound of $\frac{A_0^i}{\varepsilon}$ which is defined based on the observer gain. Therefore by changing the observer gain, the threshold value changes. The threshold values obtained by using Eq. 3.38 are constant since they are defined based on bound of the estimation error after convergence to stability region and a constant bound for the uncertainty. In contrast, the time-varying thresholds in Armaou and Demetriou (2008) and Zhang *et al.* (2010b) are due to considering the estimation error present before convergence. Note that the thresholds defined in this work can be readily made time-varying by considering the estimation error before convergence. Also in the scenario that time-varying bounds for the uncertainty are known, the level of conservatism in thresholds can be further reduced.

Remark 3.7. Note that in the case of systems with relatively low degree of nonlinearity like single link robot arm of Zhang *et al.* (2010b), the parameters in Eq. 3.38 can be specified and as a result, the Eq. 3.38 can be used directly for defining thresholds. However, when it comes to highly nonlinear systems like the CSTR example used here, it is not possible to find all of the constants in the Eq. 3.38. Instead we used simulations to determine the suprema of $\|\tilde{x}_{sub,i,j}(t_{k+1}) - x_{sub,i,j}(t_{k+1})\|$ and $\|x_{sub,i,j}(t_{k+1}) - \hat{x}_{sub,i,j}(t_{k+1})\|$, to in turn utilize as the threshold values suggested by Eq. 3.27. Note that the main purpose of deriving a mathematical formula for thresholds (Eq. 3.38) in this work is to rigorously establish the ability of the proposed scheme to achieve fault detection and isolation in the presence of uncertainty.

A fault is declared when at least one of the residuals breach their threshold i.e. $r_{i,d} > \delta_i$ for some i, and we denote this time as t_d. It follows from Lemma 3.1 that there will be no false alarms in the proposed FDI scheme. Corollary 3.1 establishes that a residual designed to be insensitive to a specific fault scenario (using the approach in Du *et al.* (2013)) will remain insensitive even in the presence of uncertainty when thresholds are picked using the proposed approach (preserving the unique breaching pattern necessary for FDI). To this end, let r^i_{ins} denote the vector of residuals designed to be insensitive to the ith fault scenario Θ_{f_i}. Corollary 3.1 establishes this property.

Corollary 3.1. *Consider the system of Eq. 3.1, for which Assumptions 3.1-3.6 hold and the fault detection and isolation framework characterized by residuals and thresholds described by Eq. 3.25 and Eq. 3.38, respectively and with* $\Theta_{f,i}(t) \neq 0$, $r_o \in r^i_{ins}$, $r_o \leq \delta_o \; \forall \; t > t_f \geq t_{k'}$ *holds.*

Proof. Consider the system of Eq. 3.1 under fault free conditions and let $r_o \in r^i_{ins}$.

From lemma 1 for $t < t_f$, we have:

$$r_o = \|\tilde{x}_{sub,o,j} - \hat{x}_{sub,o,j}\| \leq \delta_o = r^o_{pp} + r^o_{ep} \tag{3.39}$$

where

$$\|\tilde{x}_{sub,o,j} - x_{sub,o,j}\| \leq r^o_{pp} = M^o\Delta + K_{\tilde{e}} \tag{3.40}$$

and

$$\|x_{sub,o,j} - \hat{x}_{sub,o,j}\| \leq r^o_{ep} = L_{2,sub,o,j}\beta_3(\varepsilon)E_{s,sub,o,j} + \bar{\theta}_{sub,o}\Delta \tag{3.41}$$

Now we show that for $t > t_f$ $\|\tilde{x}_{sub,o,j} - x_{sub,o,j}\| \leq r^o_{pp}(\bar{\theta}_{sub}, t)$ and $\|x_{sub,o,j} - \hat{x}_{sub,o,j}\| \leq r^o_{ep}(\bar{\theta}_{sub}, t)$. Since the governing equation of scaled estimation error corresponding to r_0 after fault occurrence is same as fault free condition:

$$\dot{e^o} = \frac{1}{\varepsilon^o}A^o_0 e^o + B^o(\phi^o(\bar{u}_f) - \phi^o_0(\bar{u}_f)) + \eta_o \tag{3.42}$$

Thus, for any arbitrarily sampling time after fault occurrence, Eq. 3.41 still holds.

After fault occurrence, $\tilde{x}_{sub,o,j}$ evolves as follows:

$$\dot{\tilde{x}}_{sub,o,j} = f_{sub,o}([\tilde{x}^{\mathrm{T}}_{sub,o,j}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,o,j}]^{\mathrm{T}}) + G_{sub,o}([\tilde{x}^{\mathrm{T}}_{sub,i,j}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,o,j}]^{\mathrm{T}})\bar{u}_{f,i} \tag{3.43}$$

It follows from Eq. 3.43 that the governing equation of $\tilde{e}_{o,k} = \tilde{x}_{sub,o,j} - x_{sub,o}$ is the same as fault free condition:

$$\begin{aligned}\dot{\tilde{e}}_{o,k} = {} & f_{sub,o}([\tilde{x}^{\mathrm{T}}_{sub,o,j}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,o,j}]^{\mathrm{T}}) - f_{sub,o}([x^{\mathrm{T}}_{sub,o}, \bar{x}^{\mathrm{T}}_{sub,i}]^{\mathrm{T}}) + \\ & (G_{sub,o}([\tilde{x}^{\mathrm{T}}_{sub,o,j}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,o,j}]^{\mathrm{T}}) - G_{sub,o}([x^{\mathrm{T}}_{sub,o}, \bar{x}^{\mathrm{T}}_{sub,i}]^{\mathrm{T}}))\bar{u}_{f,i} - \theta_{sub,i}([x^{\mathrm{T}}_{sub,o}, \bar{x}^{\mathrm{T}}_{sub,o}]^{\mathrm{T}}, \bar{u}_{f,i}, t)\end{aligned} \tag{3.44}$$

And as a result, for $t > t_f$, $\|\tilde{x}_{sub,i,j} - x_{sub,i,j}\| \leq r^o_{pp}(\bar{\theta}_{sub}, t)$. This concludes the proof of Corollary 3.1. □

3.4.1 Detectability Analysis for Simultaneous Actuator and Sensor Faults

The success of the FDI scheme relies on some of the residuals not breaching thresholds, while other, dedicated residuals breaching thresholds. Having shown the first part, in this section we establish the conditions necessary for the residuals designed to be sensitive to a particular fault to breach their thresholds in the presence of uncertainty via a detectability analysis (see e.g, Zhang *et al.* (2010b), Frank (1990), Polycarpou and Trunov (2000), Chen and Patton (2012), Isermann (2006), Blanke *et al.* (2006) and Ding (2008) for similar analysis for their FDI designs). Theorem 3.2 presents the sufficient conditions for simultaneous single actuator and single sensor faults to be detectable by the proposed FDI framework. To this end, let $r_{\bar{u}_{f_i},\bar{y}_{f_i}}$ denote a sensitive residual to simultaneous faults u_{f_i} and y_{f_i} defined as:

$$r_{\bar{u}_{f_i},\bar{y}_{f_i},k+1} = \|\tilde{x}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}(t_{k+1}) - \hat{x}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}(t_{k+1})\| \tag{3.45}$$

Theorem 3.2. *Consider the system of Eq. 3.1, for which Assumptions 3.1-3.6 hold and the fault detection and isolation framework characterized by residual and threshold described by Eq. 3.25 and Eq. 3.38, respectively, and that a single actuator u_{f_i} and single sensor fault y_{f_i} occur simultaneously at time t_f: If there exists an interval of*

time $[t_f, t_d]$ *where* $t_f \geq t_{k'}$*, such that the fault functions* u_{f_i} *and* y_{f_i} *satisfy*

$$\|\frac{1}{L'_{2,\bar{u}_{f_i}}\beta_3^{\bar{u}_{f_i},\bar{y}_{f_i}}(\varepsilon_j)}\|\vartheta(\frac{A_{0,sub,\bar{u}_{f_i},\bar{y}_{f_i}}}{\varepsilon_j}, H_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}, y_{f_i})\|$$
$$+\|Dev_{t_f\ldots t_d}(\tilde{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}, \bar{u}_{f_i}, f_{d_j}, x^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}, u_{f_i})\| + \|f_{d,u_{f_i}}\| - \delta'_{\bar{u}_{f_i},\bar{y}_{f_i}} - \delta_{\bar{u}_{f_i},\bar{y}_{f_i}}\| > \delta_{\bar{u}_{f_i},\bar{y}_{f_i}} \tag{3.46}$$

where $Dev_{t_f\ldots t_d}(\tilde{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}, \bar{u}_{f,i}, f_{d_j}, x^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}, u_{f_i}) = \int_{t_f}^{t_{f+1}}(dev(\tilde{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}, \bar{u}_{f,i}, f_{d_j}) - G_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}([x^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}, \bar{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}]^{\mathrm{T}})u_{f_i})d\tau + \cdots + \int_{t_{d-1}}^{t_d}(dev(\tilde{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}, \bar{u}_{f,i}, f_{d_j}) - G_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}([x^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}, \bar{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}]^{\mathrm{T}})u_{f_i})d\tau$*,* $dev(\tilde{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}, \tilde{x}^{\mathrm{T}}_{sub,u_{f_i},\bar{y}_{f_i}}, u_{sub,i}, f_{d_j}) = \tilde{f}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}([\tilde{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}]^{\mathrm{T}}) + \tilde{G}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}([\tilde{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}]^{\mathrm{T}})u_{sub,i} - \tilde{f}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}([\tilde{x}^{\mathrm{T}}_{sub,u_{f_i},y_{f_i}}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,u_{f_i},y_{f_i}}]^{\mathrm{T}}) - \tilde{G}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}([\tilde{x}^{\mathrm{T}}_{sub,u_{f_i},y_{f_i}}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,u_{f_i},y_{f_i}}]^{\mathrm{T}})u_{sub,i}$*, where* $u_{sub,i}$ *is the subset of inputs corresponding to* $G_{sub,\bar{u}_f,\bar{y}_f}$*, where* j *refers to the corresponding observer used for defining* $r_{\bar{u}_{f_i},\bar{y}_{f_i}}$ *and* f_{d_j} *is the deviation of state estimates value from system states after fault occurrence:*

$$\begin{aligned} f_{d_j} &= \hat{x}_{\bar{u}_{f_i},\bar{y}_{f_i}} - \hat{x}_{u_{f_i},y_{f_i}} = T_j'^{-1}(\hat{\zeta}_{\bar{u}_{f_i},\bar{y}_{f_i}}, u_i) - T_j'^{-1}(\hat{\zeta}_{faultfree,\bar{u}_{f_i},\bar{y}_{f_i}}, u_{f_i}+u_i) \\ &= T_j'^{-1}(\hat{\zeta}_{\bar{u}_{f_i},\bar{y}_{f_i}}, u_i) - T_j'^{-1}(\hat{\zeta}_{\bar{u}_{f_i},\bar{y}_{f_i}}, u_{f_i}+u_i) + T_j'^{-1}(\hat{\zeta}_{\bar{u}_{f_i},\bar{y}_{f_i}}, u_{f_i}+u_i) \\ &\quad - T_j'^{-1}(\hat{\zeta}_{faultfree,\bar{u}_{f_i},\bar{y}_{f_i}}, u_{f_i}+u_i) \end{aligned} \tag{3.47}$$

where $\hat{\zeta}_{\bar{u}_{f_i},\bar{y}_{f_i}} = \hat{\zeta}_{faultfree,\bar{u}_{f_i},\bar{y}_{f_i}} + H_{\bar{u}_{f_i},\bar{y}_{f_i}}\int_{t_f}^{t_{f+1}} e^{(A_i - H_{\bar{u}_{f_i},\bar{y}_{f_i}}C_i)(t_{f+1}-\tau)}y_{f_i}d\tau + \ldots + H_{\bar{u}_{f_i},\bar{y}_{f_i}}\int_{t_{d-1}}^{t_d} e^{(A_i - H_{\bar{u}_{f_i},\bar{y}_{f_i}}C_i)(t_d-\tau)}y_{f_i}d\tau$*,* $\vartheta(\frac{A_{0,sub,i,j}}{\varepsilon_j}, H_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}, y_{f_i}) = \int_{t_f}^{t_{f+1}} e^{\frac{A_{0,sub,\bar{u}_{f_i},\bar{y}_{f_i}}}{\varepsilon_j}(t_{f+1}-\tau)}[D_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}]^{-1}H_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}y_{f_i}d\tau + \ldots + \int_{t_{d-1}}^{t_d} e^{\frac{A_{0,sub,\bar{u}_{f_i},\bar{y}_{f_i}}}{\varepsilon_j}(t_d-\tau)}[D_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}]^{-1}H_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}y_{f_i}d\tau$*,* $f_{d,u_{f_i}} = T_j^{-1}(\zeta_{\bar{u}_{f_i},\bar{u}_{f_i}}, u_{f_i}+u_i) - T_j^{-1}(\zeta_{\bar{u}_{f_i},\bar{u}_{f_i}}, u_i)$ *and* $\delta'_{\bar{u}_{f_i},\bar{y}_{f_i}} = E'_{s,\bar{u}_{f_i},\bar{y}_{f_i}} = (\frac{1}{L'_{2,u_i}\beta_3^{\bar{u}_{f_i},\bar{y}_{f_i}}(\varepsilon_j)} - L_{2,\bar{u}_{f_i}}\beta_3^{\bar{u}_{f_i},\bar{y}_{f_i}}(\varepsilon_j))E_{s,\bar{u}_{f_i},\bar{y}_{f_i}}$*, then the fault is detected, i.e.* $r_{\bar{u}_{f_i},\bar{y}_{f_i},d} > \delta_{\bar{u}_{f_i},\bar{y}_{f_i}}$*.*

Proof. After fault occurrence (i.e., $t > t_f$), the prediction model corresponding to $r_{\bar{u}_{f_i},\bar{y}_{f_i}}$ takes the following form:

$$\begin{aligned}\dot{\tilde{x}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}} &= \tilde{f}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}([\tilde{x}^{\mathrm{T}}_{sub,u_{f_i},y_{f_i}}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,u_{f_i},y_{f_i}}]^{\mathrm{T}}) \\ &\quad + \tilde{G}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}([\tilde{x}^{\mathrm{T}}_{sub,u_{f_i},y_{f_i}}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,u_{f_i},y_{f_i}}]^{\mathrm{T}})u_{sub,i} + dev(\tilde{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}, \bar{u}_{f_i}, f_{d_j})\end{aligned} \tag{3.48}$$

where $u_{sub,i}$ is properly defined and $dev(\tilde{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}, u_{sub,i}, f_{d_j}) =$
$\tilde{f}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}([\tilde{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}]^{\mathrm{T}}) + \tilde{G}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}([\tilde{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}]^{\mathrm{T}})u_{sub,i}$
$- \tilde{f}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}([\tilde{x}^{\mathrm{T}}_{sub,u_{f_i},y_{f_i}}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,u_{f_i},y_{f_i}}]^{\mathrm{T}}) - \tilde{G}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}([\tilde{x}^{\mathrm{T}}_{sub,u_{f_i},y_{f_i}}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,u_{f_i},y_{f_i}}]^{\mathrm{T}})u_{sub,i}$ and f_{d_j} is the deviation of state estimates from system states, after fault occurrence. To calculate f_{d_j}, we consider the corresponding observer to $r_{\bar{u}_{f_i},\bar{y}_{f_i}}$, and focus on the residual that breaches the threshold due to the occurrence of both the sensor and actuator fault (note that the analysis for residuals that are breached due to say, only the actuator or only the sensor faults are special cases of the below analysis). We then have:

$$\begin{aligned}\dot{\hat{\zeta}}_{\bar{u}_{f_i},\bar{y}_{f_i}} &= A_i\hat{\zeta}_{\bar{u}_{f_i},\bar{y}_{f_i}} + B_i\phi_0 + H_{\bar{u}_{f_i},\bar{y}_{f_i}}(C_i\zeta_{\bar{u}_{f_i},\bar{y}_{f_i}} + y_{f_i} - C_i\hat{\zeta}_{\bar{u}_{f_i},\bar{y}_{f_i}}) \\ \hat{\zeta}_{\bar{u}_{f_i},\bar{y}_{f_i}}(t_k) &= T'(\hat{x}_i(t_k), u_i(t_k))\end{aligned} \tag{3.49}$$

By integration from t_f to t_d, we get:

$$\begin{aligned}\hat{\zeta}_{\bar{u}_{f_i},\bar{y}_{f_i}} &= \hat{\zeta}_{faultfree,\bar{u}_{f_i},\bar{y}_{f_i}} + H_{\bar{u}_{f_i},\bar{y}_{f_i}}\int_{t_f}^{t_{f+1}} e^{(A_i - H_{\bar{u}_{f_i},\bar{y}_{f_i}}C_i)(t_{f+1}-\tau)}y_{f_i}d\tau + \dots \\ &\quad + H_{\bar{u}_{f_i},\bar{y}_{f_i}}\int_{t_{d-1}}^{t_d} e^{(A_i - H_{\bar{u}_{f_i},\bar{y}_{f_i}}C_i)(t_d-\tau)}y_{f_i}d\tau\end{aligned} \tag{3.50}$$

where $\hat{\zeta}_{faultfree,\bar{u}_{f_i},\bar{y}_{f_i}} = \hat{\zeta}_{\bar{u}_{f_i},\bar{y}_{f_i}}(t_f) + \int_{t_f}^{t_{f+1}} e^{(A_i - H_{\bar{u}_{f_i},\bar{y}_{f_i}}C_i)(t_{f+1}-\tau)}(B_i\phi_0 + H_{\bar{u}_{f_i},\bar{y}_{f_i}}C_i\zeta_{\bar{u}_{f_i},\bar{y}_{f_i}})d\tau$

$+\cdots+\int_{t_{d-1}}^{t_d} e^{(A_i-H_{\bar{u}_{f_i},\bar{y}_{f_i}}C_i)(t_d-\tau)}(B_i\phi_0+H_{\bar{u}_{f_i},\bar{y}_{f_i}}C_i\zeta_{\bar{u}_{f_i},\bar{y}_{f_i}})d\tau$. Therefore, f_{d_j} is formulated as below:

$$\begin{aligned} f_{d_j} &= \hat{x}_{\bar{u}_{f_i},\bar{y}_{f_i}} - \hat{x}_{u_{f_i},y_{f_i}} = T_j'^{-1}(\hat{\zeta}_{\bar{u}_{f_i},\bar{y}_{f_i}},u_i) - T_j'^{-1}(\hat{\zeta}_{faultfree,\bar{u}_{f_i},\bar{y}_{f_i}},u_{f_i}+u_i) \\ &= T_j'^{-1}(\hat{\zeta}_{\bar{u}_{f_i},\bar{y}_{f_i}},u_i) - T_j'^{-1}(\hat{\zeta}_{\bar{u}_{f_i},\bar{y}_{f_i}},u_{f_i}+u_i) + T_j'^{-1}(\hat{\zeta}_{\bar{u}_{f_i},\bar{y}_{f_i}},u_{f_i}+u_i) \\ &- T_j'^{-1}(\hat{\zeta}_{faultfree,\bar{u}_{f_i},\bar{y}_{f_i}},u_{f_i}+u_i) \end{aligned} \tag{3.51}$$

Substituting f_{d_j} into Eq. 3.31 gives:

$$\begin{aligned} \dot{\tilde{e}}_{\bar{u}_{f_i},\bar{y}_{f_i},k} &= \tilde{f}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}} - f_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}([x^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}},\bar{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}]^{\mathrm{T}}) + (\tilde{G}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}} \\ &- G_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}([x^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}},\bar{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}]^{\mathrm{T}})u_{sub,i} + dev(\tilde{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}},\bar{u}_{f,i},f_d) \\ &- G_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}([x^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}},\bar{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}]^{\mathrm{T}})u_{f_i} \end{aligned} \tag{3.52}$$

By integration from t_f to t_d, we obtain:

$$\tilde{e}_{\bar{u}_{f_i},\bar{y}_{f_i},t_d} = \tilde{e}_{\bar{u}_{f_i},\bar{y}_{f_i},faultfree} + Dev_{t_f\ldots t_d}(\tilde{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}},\bar{u}_{f,i,sub},f_{d_j},x^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}},u_{f_i}) \tag{3.53}$$

where $\tilde{e}_{\bar{u}_{f_i},\bar{y}_{f_i},faultfree} = \tilde{e}_{\bar{u}_{f_i},\bar{y}_{f_i},t_f} + \int_{t_f}^{t_{f+1}}(\tilde{f}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}} - f_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}([x^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}},\bar{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}]^{\mathrm{T}}) +$
$(\tilde{G}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}} - G_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}([x^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}},\bar{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}]^{\mathrm{T}}))u_i)d\tau + \ldots$
$+\int_{t_{d-1}}^{t_d}(\tilde{f}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}} - f_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}([x^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}},\bar{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}]^{\mathrm{T}})$
$+(\tilde{G}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}} - G_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}([x^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}},\bar{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}]^{\mathrm{T}}),\ u_i)d\tau,$
$Dev_{t_f\ldots t_d}(\tilde{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}},\bar{u}_{f,i,sub},f_{d_j},x^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}},u_{f_i}) = \int_{t_f}^{t_{f+1}}(dev(\tilde{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}},\bar{u}_{f,i,sub},f_{d_j})$
$-G_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}([x^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}},\bar{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}]^{\mathrm{T}})u_{f_i})d\tau + \cdots + \int_{t_{d-1}}^{t_d}(dev(\tilde{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}},\bar{u}_{f,i,sub},f_{d_j})$
$-G_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}([x^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}},\bar{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}]^{\mathrm{T}})u_{f_i})d\tau.$

Now by applying the triangular inequality, we get:

$$\begin{aligned}\|\tilde{x}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}(t_d) - x_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}(t_d)\| &= \|\tilde{e}_{\bar{u}_{f_i},\bar{y}_{f_i},k}(t_d)\| \\ &\geq \|Dev_{t_f\ldots t_d}(\tilde{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}, \bar{u}_{f,i,sub}, f_{d_j}, x^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}, u_{f_i})\| \\ &- E_p^{\bar{u}_{f_i},\bar{y}_{f_i}}\end{aligned} \tag{3.54}$$

After fault occurrence the Eq. 3.32 takes the following form:

$$\begin{aligned}\dot{e}_{sub,\bar{u}_{f_i},\bar{y}_{f_i},j} = \frac{1}{\varepsilon_j} A_{0,sub,\bar{u}_{f_i},\bar{y}_{f_i}} e_{sub,\bar{u}_{f_i},\bar{y}_{f_i}} + B_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}(\phi_{sub,\bar{u}_{f_i},\bar{y}_{f_i}} - \phi_{0,sub,\bar{u}_{f_i},\bar{y}_{f_i}}) \\ + \eta_{i,sub,\bar{u}_{f_i},\bar{y}_{f_i}} + [D_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}]^{-1} H_{sub,\bar{u}_{f_i},\bar{y}_{f_i}} y_{f_i}\end{aligned} \tag{3.55}$$

By integration from t_f to t_d, we obtain:

$$e_{sub,\bar{u}_{f_i},\bar{y}_{f_i},j}(t_d) = e_{\bar{u}_{f_i},\bar{y}_{f_i},j,faultfree} + \vartheta(\frac{A_{0,sub,\bar{u}_{f_i},\bar{y}_{f_i}}}{\varepsilon_j}, H_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}, y_{f_i}) \tag{3.56}$$

where $e_{\bar{u}_{f_i},\bar{y}_{f_i},j,faultfree} = e_{sub,\bar{u}_{f_i},\bar{y}_{f_i},j}(t_f) + \int_{t_f}^{t_{f+1}} e^{\frac{A_{0,sub,\bar{u}_{f_i},\bar{y}_{f_i}}}{\varepsilon_j}(t_{f+1}-\tau)}(B_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}(\phi_{sub,\bar{u}_{f_i},\bar{y}_{f_i}} - \phi_{0,sub,i,j}) + \eta_{i,sub,\bar{u}_{f_i},\bar{y}_{f_i}})d\tau + \cdots + \int_{t_{d-1}}^{t_d} e^{\frac{A_{0,sub,\bar{u}_{f_i},\bar{y}_{f_i}}}{\varepsilon_j}(t_d-\tau)}(B_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}(\phi_{sub,\bar{u}_{f_i},\bar{y}_{f_i}} - \phi_{0,sub,\bar{u}_{f_i},\bar{y}_{f_i}}) + \eta_{i,sub,\bar{u}_{f_i},\bar{y}_{f_i}})d\tau$,
$\vartheta(\frac{A_{0,sub,\bar{u}_{f_i},\bar{y}_{f_i}}}{\varepsilon_j}, H_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}, y_{f_i}) = \int_{t_f}^{t_{f+1}} e^{\frac{A_{0,sub,\bar{u}_{f_i},\bar{y}_{f_i}}}{\varepsilon_j}(t_{f+1}-\tau)}[D_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}]^{-1} H_{sub,\bar{u}_{f_i},\bar{y}_{f_i}} y_{f_i} d\tau + \cdots + \int_{t_{d-1}}^{t_d} e^{\frac{A_{0,sub,\bar{u}_{f_i},\bar{y}_{f_i}}}{\varepsilon_j}(t_d-\tau)}[D_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}]^{-1} H_{sub,\bar{u}_{f_i},\bar{y}_{f_i}} y_{f_i} d\tau$. By applying the triangular inequality, we get:

$$\|e_{\bar{u}_{f_i},\bar{y}_{f_i}}(t_{k+1})\| \geq E_{s,\bar{u}_{f_i},\bar{y}_{f_i}} - \|\vartheta(\frac{A_{0,sub,\bar{u}_{f_i},\bar{y}_{f_i}}}{\varepsilon_j}, H_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}, y_{f_i})\| \tag{3.57}$$

For the residual that breaches the threshold due to both sensor and actuator fault,

we have:

$$\begin{aligned} x_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}(t_d) - \hat{x}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}(t_d) &= T_j^{-1}(\zeta_{\bar{u}_{f_i},\bar{u}_{f_i}}, u_{f_i} + u_i) - T_j'^{-1}(\hat{\zeta}_{\bar{u}_{f_i},\bar{u}_{f_i}}, u_i) \\ &= T_j^{-1}(\zeta_{\bar{u}_{f_i},\bar{u}_{f_i}}, u_{f_i} + u_i) - T_j^{-1}(\zeta_{\bar{u}_{f_i},\bar{u}_{f_i}}, u_i) \\ &+ T_j^{-1}(\zeta_{\bar{u}_{f_i},\bar{u}_{f_i}}, u_i) - T_j'^{-1}(\hat{\zeta}_{\bar{u}_{f_i},\bar{u}_{f_i}}, u_i) \end{aligned} \tag{3.58}$$

By using Lipschitz property of $T_j'^{-1}$ and Eq. 3.18, after applying the triangular inequality, we get:

$$\begin{aligned} &\|x_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}(t_d) - \hat{x}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}(t_d)\| \geq L_{2,u_i}\beta_3(\varepsilon_j)E_{s,\bar{u}_{f_i},\bar{y}_{f_i}} + \bar{\theta}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}\Delta + E'_{s,\bar{u}_{f_i},\bar{y}_{f_i}} \\ &+ \|f_{d,u_{f_i}}\| - \frac{1}{L'_{2,u_i}\beta_3^{\bar{u}_{f_i},\bar{y}_{f_i}}(\varepsilon_j)}\|\vartheta(\frac{A_{0,sub,\bar{u}_{f_i},\bar{y}_{f_i}}}{\varepsilon_j}, H_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}, y_{f_i})\| \end{aligned} \tag{3.59}$$

where $E'_{s,\bar{u}_{f_i},\bar{y}_{f_i}} = (\frac{1}{L'_{2,u_i}\beta_3^{\bar{u}_{f_i},\bar{y}_{f_i}}(\varepsilon_j)} - L_{2,u_i}\beta_3^{\bar{u}_{f_i},\bar{y}_{f_i}}(\varepsilon_j))E_{s,\bar{u}_{f_i},\bar{y}_{f_i}}$ and $f_{d,u_{f_i}} = T_j^{-1}(\zeta_{\bar{u}_{f_i},\bar{u}_{f_i}}, u_{f_i} + u_i) - T_j^{-1}(\zeta_{\bar{u}_{f_i},\bar{u}_{f_i}}, u_i)$. Therefore by using Eqs. 3.54, 3.59 and triangular inequality, we obtain:

$$\begin{aligned} &r_{i,d} = \|\tilde{x}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}(t_d) - \hat{x}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}(t_d)\| \geq \|\frac{1}{L'_{2,u_i}\beta_3^{\bar{u}_{f_i},\bar{y}_{f_i}}(\varepsilon_j)}\|\vartheta(\frac{A_{0,sub,\bar{u}_{f_i},\bar{y}_{f_i}}}{\varepsilon_j}, \\ &H_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}, y_{f_i})\| + \|f_{d,u_{f_i}}\| + \|Dev_{t_f \ldots t_d}(\tilde{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}, \bar{u}_{f_i}, f_{d_j}, x^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}, u_{f_i})\| \\ &- \delta_{\bar{u}_{f_i},\bar{y}_{f_i}} - \delta'_{\bar{u}_{f_i},\bar{y}_{f_i}}\| \end{aligned} \tag{3.60}$$

where $\delta'_{\bar{u}_{f_i},\bar{y}_{f_i}} = E'_{s,\bar{u}_{f_i},\bar{y}_{f_i}}$. Therefore it follows that if Eq. 3.46 holds, $t_f \geq t_{k'}$, $r_{\bar{u}_{f_i},\bar{y}_{f_i},d} > \delta_{i,\bar{u}_{f_i},\bar{y}_{f_i}}$. □

3.4.2 Isolability Condition

Having presented the detectability condition corresponding to different faulty scenarios, Theorem 3.3 presents the fault isolation logic for the identification of faulty component that also serves as the FDI mechanism. The proof of Theorem 3.3 follows along similar lines as Theorem 1 in Du *et al.* (2013), and is omitted here.

Theorem 3.3. *Consider the system of Eq. 3.1, for which Assumptions 3.1-3.6 hold. If* $t \geq t_d$ *and* $r_{i,t_d} > \delta_i$ *for all* $i \in \{1, \ldots, n_f\} \backslash j$*, then* $\Theta_{f,j}(t) \neq 0$ *for some* $t \in [t_d, t_{d+1})$.

Remark 3.8. Theorem 3.3 states that the proposed FDI framework will result in a unique breaching pattern for every single fault scenario, leading to FDI (see e.g., Du *et al.* (2013) for further details regarding expected breaching patterns corresponding to single or multiple fault scenarios). The key when dealing with uncertainty, however, is to ensure that such unique fault signatures continue to exist in the presence of uncertainty, which is achieved by appropriate choice of the thresholds that are cognizant of the presence of uncertainty, as done in the present work.

Remark 3.9. Note that while the proposed framework explicitly considers control actuator faults, it can readily deal with other actuator faults (that are not being utilized in feedback control). In these instance, if these variables are being prescribed certain values (constant or otherwise), the proposed method can be directly utilized to detect and isolate faults in these variables. If the variable values are not measured, then these are naturally incorporated in the uncertainty term.

Remark 3.10. The proposed approach provides the specific observer and FDI design, and backs it up with a rigorous analysis that establishes error bounds and thresholds based on the system characteristics. In practice, the rigorous analysis can be utilized

Table 3.1: Faults to which the residuals are insensitive and thresholds for the fault isolation design of the example in Section 3.5 based on the proposed framework in 3.4.

Residual	Faults	Threshold	Residual	Faults	Threshold
r_1	y_{f_1}	0.1	r_2	y_{f_2}	0.1
r_3	y_{f_3}	0.1	r_4	y_{f_4}	0.1
r_5	y_{f_5}	0.5	r_6	y_{f_1}, y_{f_2}	0.1
r_7	y_{f_1}, y_{f_3}	0.1	r_8	y_{f_1}, y_{f_4}	0.1
r_9	y_{f_1}, y_{f_5}	0.48	r_{10}	y_{f_2}, y_{f_3}	0.1
r_{11}	y_{f_2}, y_{f_4}	0.1	r_{12}	y_{f_2}, y_{f_5}	0.33
r_{13}	y_{f_3}, y_{f_4}	0.1	r_{14}	y_{f_3}, y_{f_5}	2.8
r_{15}	y_{f_4}, y_{f_5}	1.7	r_{16}	u_{f_2}	0.002
r_{17}	u_{f_2}, y_{f_5}	0.002	r_{18}	u_{f_1}	0.1
r_{19}	u_{f_1}, y_{f_1}	0.3	r_{20}	u_{f_1}, u_{f_2}	0.002
r_{21}	u_{f_1}, y_{f_2}	0.12	r_{22}	u_{f_1}, y_{f_3}	0.13
r_{23}	u_{f_1}, y_{f_4}	0.14	r_{24}	u_{f_1}, y_{f_5}	0.5
r_{25}	u_{f_2}, y_{f_1}	0.002	r_{26}	u_{f_2}, y_{f_2}	0.002
r_{27}	u_{f_2}, y_{f_3}	0.003	r_{28}	u_{f_2}, y_{f_4}	0.004

to provide confidence in the FDI capabilities of the proposed filters and to guide the selection of the parameters (as is done in the simulation results).

3.5 Simulation Example

In this section, we consider a continuous-stirred tank reactor (CSTR) example (see Du *et al.* (2013) for details regarding process model, process parameters and control design).

The process is subject to modeling uncertainty and measurement noise. In particular, the actual values of the reaction rate constants k_{1A0} and k_{2A0} are 10% less than the values in the model used for FDI. Furthermore, the flow rate fluctuates with time, with the actual flow rate being $1 + 0.05\sin(t)$ times of its nominal value. The

known bounds on each uncertainty is 15%, 15% and 5% of the absolute nominal values. The concentration and temperature measurements have combinations of 5 Hz sinusoidal noises. The magnitudes of the measurement noise over each 0.5 min follow a normal distribution with the standard deviations being 0.02 kmol/m^3 and 0.5 K for concentrations and temperatures, respectively. The noisy measurements are passed through a first-order low-pass filter with the filter time constant being 3 seconds.

Note that Assumption 3.1 holds for the simulation case study, since of all of the terms in f, g and θ are differentiable and their differentiation is continuous and $f(0) = 0$. Also, since the desired set point is met in the presence of uncertainty and in the absence of fault (see Figure 1 in Du *et al.* (2013)), by using the converse Lyapunov theorem (see e.g., Khalil (2002)), it is guaranteed that a robust Lyapunov function exists that satisfies Assumption 3.2.

Fifteen observers are designed using subsets of the available measurements to be utilized in the residual generation. Twenty-eight residuals are generated using the methodology described in Section 3.4. Among these observers, five are defined using four of the available measurements that are utilized for generation of residuals insensitive to the single sensor faults in one of the outputs or simultaneous single sensor and single actuator faults. This includes residuals r_1 to r_5 (corresponding residuals to single sensor faults) and residuals r_{17}, r_{19} and r_{21} to r_{28} (residuals corresponding to simultaneous single sensor and single actuator faults). The rest of the observers are defined using three of the available measurements that are utilized for generation of residuals insensitive to multiple sensor faults, residuals r_6 to r_{15}. Note the residuals corresponding to single and multiple actuator faults (r_{16}, r_{18} and r_{20}) are generated

using the state estimates provided by the observer that are not affected by the corresponding input(s). Note that Assumptions 3.3 and 3.6 hold for all of the designed observers. As an example, consider the case where only measurements of $y_1 = C_A$, $y_2 = C_B$ and $y_3 = C_C$ are used for observer design. For this case, one of the possible transformations is presented below:

$$T = T' + T_\theta, T' = \begin{bmatrix} C_A \\ \dot{C_A} \\ C_B \\ \dot{C_B} \\ C_C \\ \dot{C_C} \end{bmatrix}, T_\theta = \begin{bmatrix} 0 \\ \frac{0.05Fsin(t)}{V}(C_{\mathrm{A0}} - C_{\mathrm{A}}) - 0.1r_{\mathrm{1A,\ forward}} - 0.1r_{\mathrm{2A,\ forward}} \\ 0 \\ \frac{0.05Fsin(t)}{V}(C_{\mathrm{B0}} - C_{\mathrm{B}}) - 0.1r_{\mathrm{1A,\ forward}} \\ 0 \\ -\frac{0.05Fsin(t)}{V}C_{\mathrm{C}}) + 0.1r_{\mathrm{1A,\ forward}} - 0.1r_{\mathrm{2A,\ forward}} \end{bmatrix}$$

For the rest of designed observers the Assumption 5 can be verified in the same manner. Also, here we consider $\phi_0 = 0$ that is always bounded, thus Assumption 4 holds as well.

The thresholds are selected based on Eq. 3.27 via simulations. To this end, a value slightly larger than the summation of the maximum observed values for $\|\tilde{x}_{sub,i,j}(t_{k+1}) - x_{sub,i,j}(t_{k+1})\|$ and $\|x_{sub,i,j}(t_{k+1}) - \hat{x}_{sub,i,j}(t_{k+1})\|$ when the system states enters the stability region, by considering all possible combinations of the bounds on uncertainties is selected as the corresponding threshold for each residual, as shown in Table 3.1.

Remark 3.11. The proposed approach does not make any assumptions on the nature of the controller. In particular, even if the controller is able to reject the effect of the fault, the FDI mechanism enables fault detection and isolation subject to the

corresponding detectability conditions being satisfied. To check the detectability condition presented in Theorem 2, we simply calculate the infimum of $\|\|x_{sub,i,j}(t_d) - \hat{x}_{sub,i,j}(t_d)\| - \|\tilde{x}_{sub,i,j}(t_d) - x_{sub,i}(t_d)\|\|$ for each residual and we call it the detectability constant, $\bar{\delta}$. If the detectability constant is more than the value of its corresponding threshold, then the residual is expected to breach the threshold. As an example, for the case of small abrupt fault in u_{f_2} presented in the Section 5, since the detectability constants (see Table 3.2) are less than the threshold values for all of the residuals, therefore we expect that none of the residual breaches its threshold and the fault can not be detected. This is verified by the simulations.

Table 3.2: Detectability constants ($\bar{\delta}$) for each residual for a case where abrupt fault of $u_{f_2} = 0.1$ in u_2 takes place at time $t_f = 7.5$ min

Residual	$\bar{\delta}$	Threshold	Residual	$\bar{\delta}$	Threshold
r_1	0.047	0.1	r_2	0.056	0.1
r_3	0.052	0.1	r_4	0.048	0.1
r_5	0.22	0.5	r_6	0.051	0.1
r_7	0.051	0.1	r_8	0.051	0.1
r_9	0.18	0.48	r_{10}	0.05	0.1
r_{11}	0.05	0.1	r_{12}	0.33	0.33
r_{13}	0.05	0.1	r_{14}	0.824	1.7
r_{15}	0.555	2.8	r_{16}	0	0.002
r_{17}	0	0.002	r_{18}	0.059	0.1
r_{19}	0.059	0.3	r_{20}	0	0.002
r_{21}	0.0745	0.12	r_{22}	0.066	0.13
r_{23}	0.06	0.14	r_{24}	0.22	0.5
r_{25}	0	0.002	r_{26}	0	0.002
r_{27}	0	0.003	r_{28}	0	0.004

We next consider a case where abrupt faults $u_{f_2} = 0.1$ in $u_1 = C_{\mathrm{A0}}$ and $y_{f_1} = 0.1$ in $y_1 = C_{\mathrm{A}}$ (one actuator fault and one sensor fault) take place at time $t_f = 7.5$ min. The evolution of residual profile is shown in Figure 3.2. In this case the expected fault

signature is breaching of all the residuals except r_{19}. Using the threshold designed in Du *et al.* (2013), some of the residuals breach their thresholds, wherein the fault is successfully detected but is not isolated since the residual breaching profiles do not match any of the expected breaching patterns presented in Du *et al.* (2013). In particular, while we expect only r_{19} be insensitive to the fault in u_1 and y_1, the residuals r_4, r_6, r_7, r_8, r_9, r_{10} and r_{11} are alway below the thresholds suggested in Du *et al.* (2013). However, using the proposed threshold selection, only r_{19} does not breach its threshold, matching the expected unique breaching pattern, resulting in successful fault isolation (magnified version of evolution of residual profiles for r_7 and r_{19} are shown in Figure 3.3).

Remark 3.12. Note that while residuals are defined using the same methodology as proposed in Du *et al.* (2013), definition of thresholds is the key difference between the proposed scheme in this work and Du *et al.* (2013). In Du *et al.* (2013), thresholds are selected by using normal operating data for residuals plus some additional positive value to avoid possible false alarms due to plant model mismatch and uncertainty. This results in conservatively large values for thresholds and increased number of missed faults, along with the inability to achieve FDI (as demonstrated by the simulation example). However, by selecting thresholds as suggested by Eq. 3.27, the number of missed faults is reduced by explicitly accounting for the presence of uncertainty in the design. This is achieved by selecting the smallest possible values for thresholds while still guaranteeing that no false alarm before fault occurrence are triggered (see Figure 3.2 for an illustration). Thus, using thresholds values suggested in Du *et al.* (2013), the FDI scheme is not able to isolate the fault since the thresholds are selected conservatively large to avoid any false alarm, resulting in missing

the location of the fault. In contrast, using the threshold values suggested in Table 3.1, the fault is successfully isolated.

We next consider a case where incipient faults of $u_{f_2} = (5 + 0.1 \sin t)(1 - e^{t_f - t})$ in $u_2 = T_0$ and $y_{f_2} = (0.2 + 0.2 \sin t)(1 - e^{t_f - t})$ in $y_2 = C_{\mathrm{B}}$ (one actuator fault and one sensor fault) take place at time $t_f = 7.5$ min. The evolution of residual profile is shown in Figure 3.4. The expected unique fault signature in this case is breaching of all of the residuals except r_{26}. Since all of the residuals breach their thresholds except r_{26}, which is designed to be insensitive to u_{f2} and y_{f2} (see Table 3.1), faults in u_2 and y_2 are isolated.

3.6 Conclusions

In this work, we addressed the problem of actuator and sensor fault detection and isolation of control affine nonlinear systems subject to uncertainty. An FDI framework was proposed and fault detectability and isolability conditions were rigorously derived. Finally, the efficacy of the fault isolation framework subject to uncertainty and measurement noise was illustrated using a chemical reactor example.

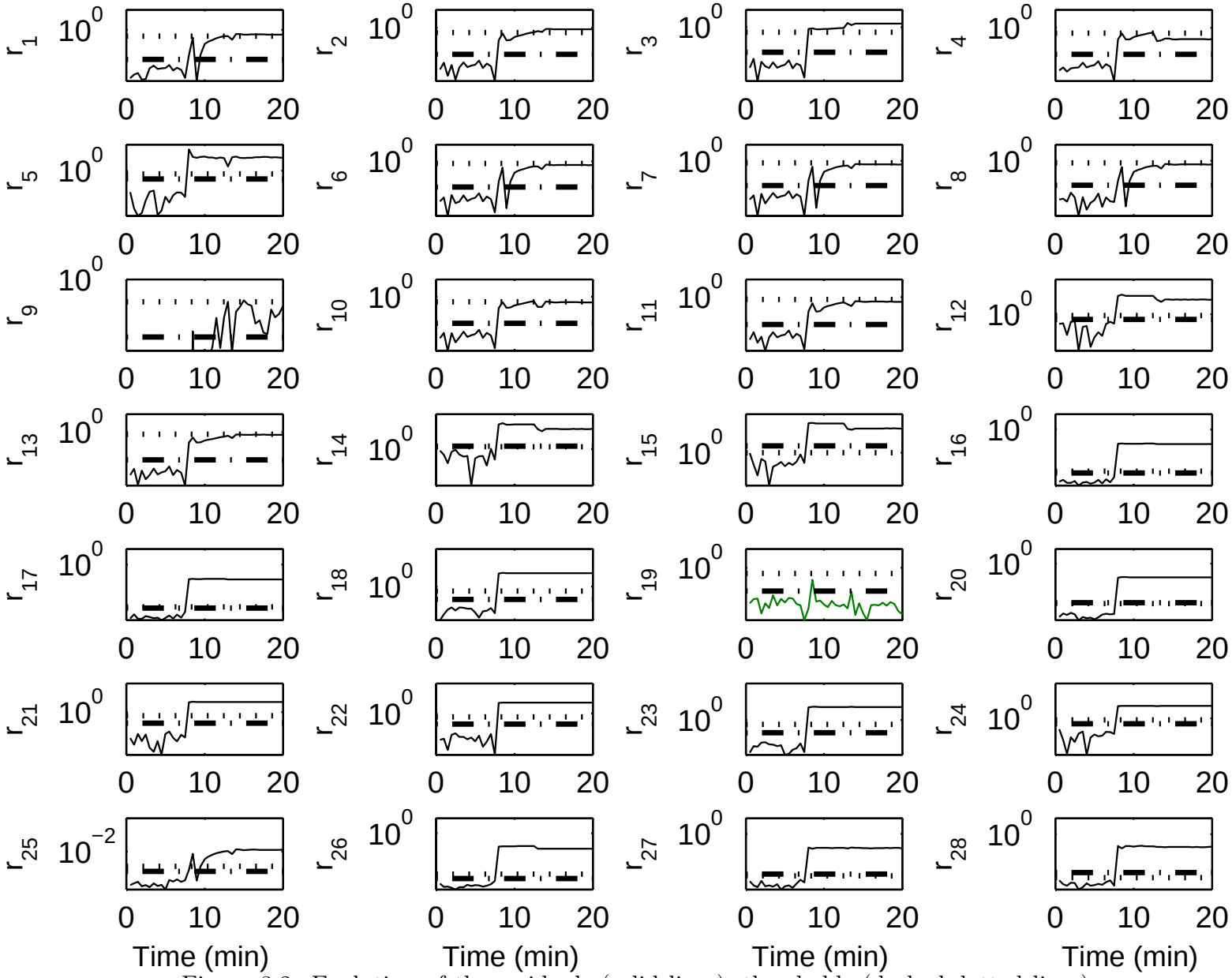


Figure 3.2: Evolution of the residuals (solid lines), thresholds (dashed-dotted lines) and thresholds designed in Du *et al.* (2013) (dotted lines). Using the thresholds proposed in Du *et al.* (2013), the residuals do not follow any of expected breaching patterns which results only in fault detection. By utilizing the thresholds designed in this work, all of the residuals breach their thresholds except for r_{19}. This corresponds to fault signature of fault only in u_1 and y_1 and as a result, faults in u_1 and y_1 are isolated.

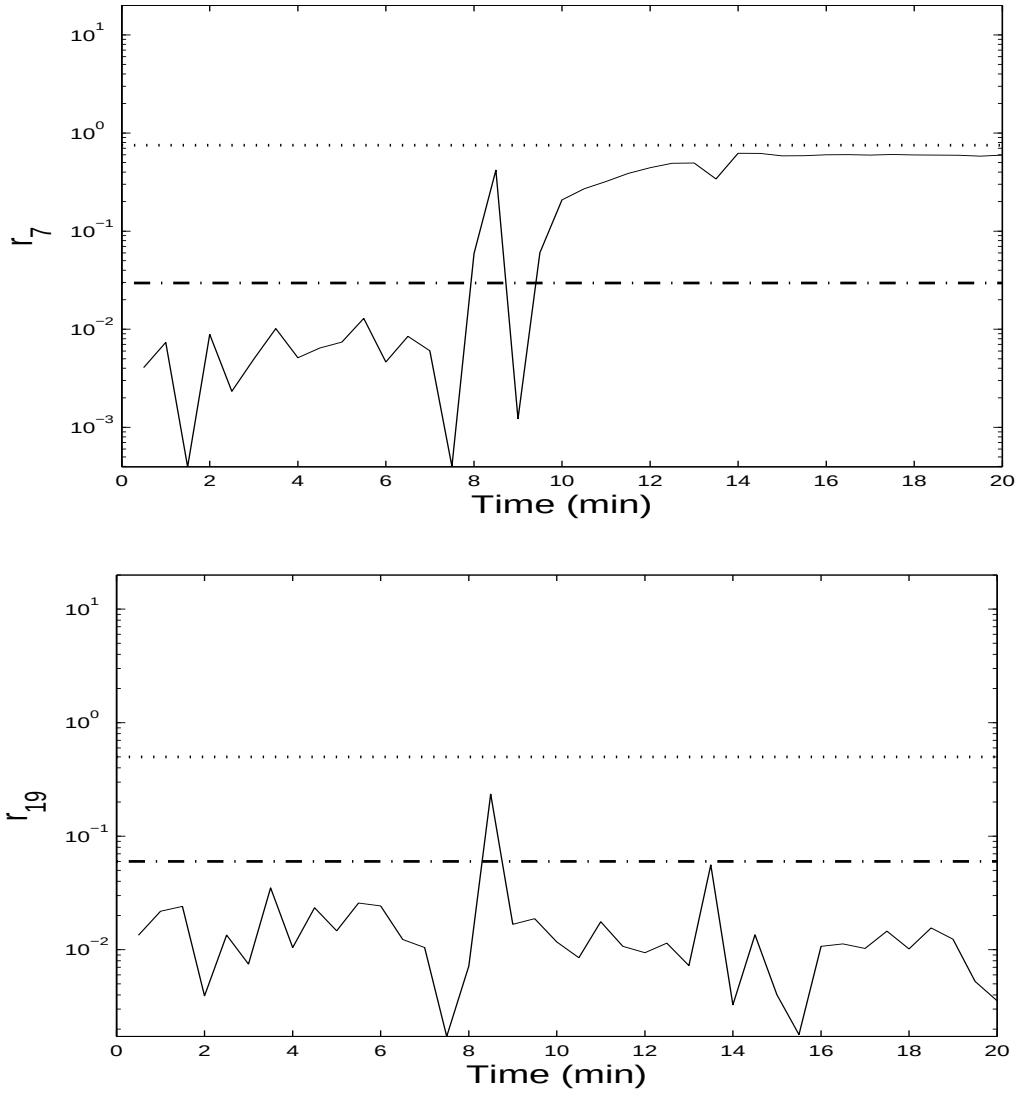


Figure 3.3: Evolution of the residuals for r_7 and r_{19} (solid lines), thresholds (dashed-dotted lines) and thresholds designed in Du *et al.* (2013) (dotted lines). Using the thresholds proposed in Du *et al.* (2013), both of the residual r_7 and r_{19} do not breach their thresholds which results only in fault detection. By utilizing the thresholds designed in this work, only r_{19} does not breach its threshold, matching a unique fault signature (see Table 3.1), and as a result, the fault scenario is isolated.

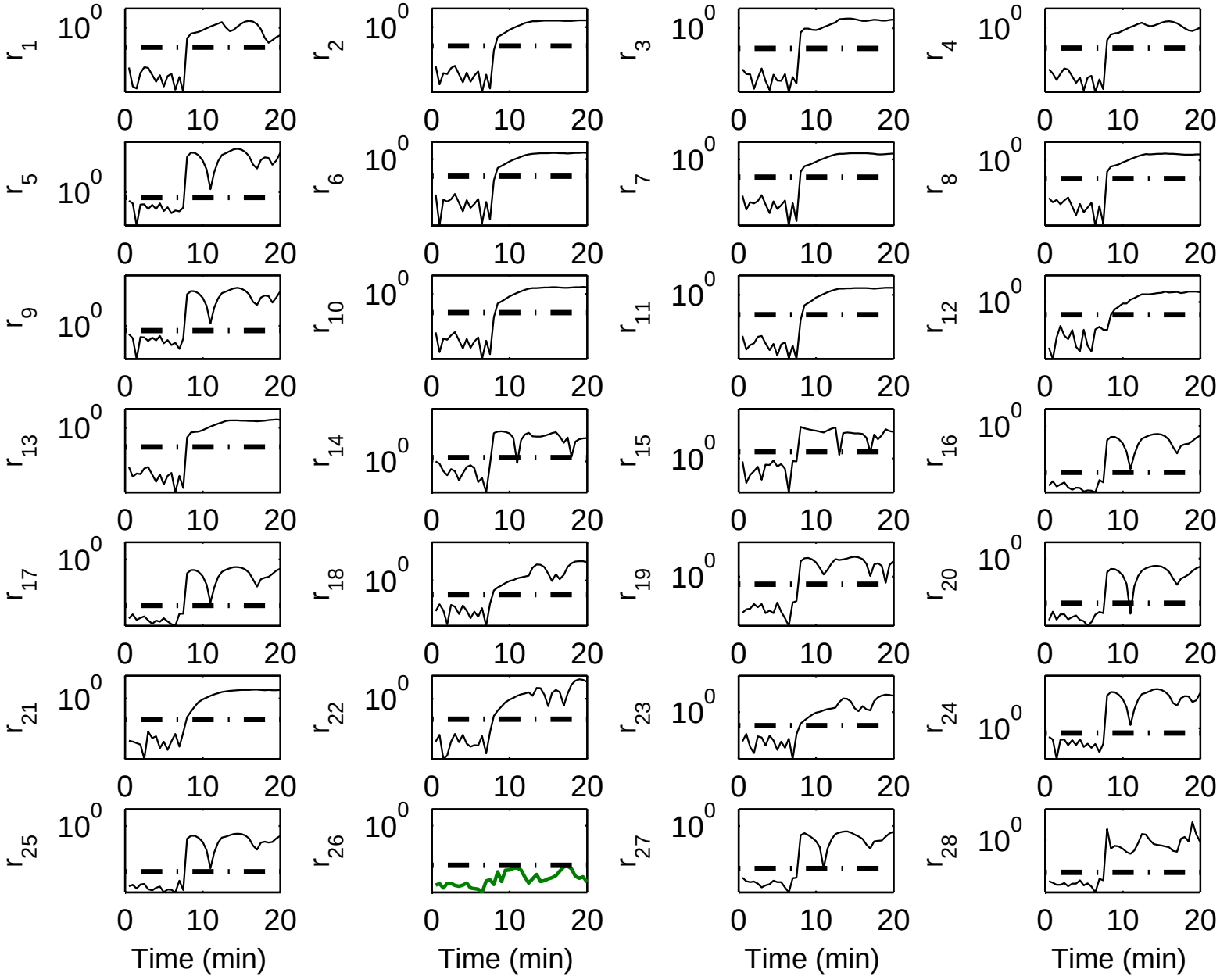


Figure 3.4: Evolution of the residuals (solid lines), thresholds (dashed-dotted lines). All the residuals breach their thresholds except for r_{26} matching the unique fault signature for a fault in u_2 and y_2 leading to fault isolation.

Chapter 4

Distributed fault diagnosis for networked nonlinear uncertain systems

The contributions of this chapter have been submitted for publication in:

Journal Papers:

Shahnazari, H. and Mhaskar, P. Distributed fault diagnosis for networked nonlinear uncertain systems, *Computers & Chemical Engineering*, submitted.

4.1 Introduction

In the previous chapter, the problem of simultaneous actuator and sensor fault diagnosis was addressed while explicitly accounting for process nonlinearities and uncertainties as some of the complexities existing in modern dynamical systems. The complexity also can be due to high dimensionality with strong interconnections between subsystems leading to their recognition as networked systems, and the implementation of extensive automation strategies (see e.g., Yan and Edwards (2008)). Extensive use of automation, however, coupled with the networked nature of the systems, also makes the major control equipments such as actuators and sensors more susceptible to faults. In particular, the faults can propagate in the network and cause major failures, necessitating fault-detection and isolation (FDI) strategies. This chapter is dedicated to fault diagnosis design for networked systems.

Where possible, designing a centralized FDI scheme is the most intuitive solution for this problem. There exist a plethora of research in designing centralized fault detection and isolation frameworks for monitoring of engineering systems (see e.g., Frank (1990); De Persis and Isidori (2001); Yan and Edwards (2007); Mhaskar *et al.* (2008); Zhang *et al.* (2010b); Du and Mhaskar (2014); Du *et al.* (2013); Shahnazari and Mhaskar (2016)). However, a centralized FDI scheme may fail or be unimplementable due to lack of existence of enough centralized computational capabilities or communication infrastructure. More importantly, there are reliability issues associated with utilizing a centralized FDI scheme for a networked system, since a failure in the FDI scheme with centralized architecture can lead to interruption in monitoring of all of the subsystems. When the interconnections are weak, using a decentralized FDI scheme composed of independent LFDI schemes for each subsystem can be a solution

to this problem (see e.g., Yan and Edwards (2008); Du *et al.* (2011)). However, when the interconnections are not negligible, utilizing a decentralized architecture for FDI design can result in false alarms or missed faults in LFDI schemes. In this case, an alternative approach can be designing an FDI scheme with a distributed architecture. In the distributed architecture, a LFDI scheme is designed for each subsystem while the LFDI schemes can communicate to exchange information.

There are some results that only consider problem of fault detection in networked and large scale systems using a distributed architecture (see e.g., Boem *et al.* (2017)). In Keliris *et al.* (2015), an integrated distributed fault detection scheme is proposed for detection of sensor and process faults in nonlinear uncertain discrete systems. However, fault isolation is not achieved. In Zhang and Zhang (2012), a distributed actuators FDI scheme is proposed for a class of interconnected uncertain nonlinear systems using adaptive estimation techniques. In Ferrari *et al.* (2012), a distributed framework is presented for diagnosing single actuator and single process faults in nonlinear uncertain large-scale discrete-time systems using an adaptive approximation based technique. In Peng *et al.* (2015), a distributed data based actuator fault identification scheme is presented for linear networked process systems. In Reppa *et al.* (2015), a distributed sensor fault diagnosis for a network of interconnected cyber-physical system (CPS)s presented. In Yin and Liu (2017), a distributed FDI scheme is proposed for cascade networked systems in the absence of uncertainty. However, none of the existing results have addressed the problem of isolation of multiple actuator faults or simultaneous actuator and sensor faults in uncertain nonlinear networked systems. Also, the results addressing isolation of actuator faults are based on the assumption that full state measurements are available (see e.g., Ferrari *et al.*

(2012)) or only valid for cascade process networks in the absence of uncertainty (see e.g., Yin and Liu (2017)). More importantly, the conditions under which the network structure allows fault isolation in the presence of uncertainty is not discussed in the existing results available in the literature (see e.g., Zhang and Zhang (2012), Ferrari *et al.* (2012), Keliris *et al.* (2015) and Reppa *et al.* (2015)).

Motivated by the above considerations, this chapter addresses the problem of simultaneous fault diagnosis in nonlinear uncertain networked systems. The rest of this chapter is organized as follows: In Section 4.2, the system description is presented. Then, in preparation of subsequent FDI filters that utilize state estimators, boundedness of estimation error in the presence of uncertainty and exchange of information is established in Section 4.3. Next, a distributed fault diagnosis framework is presented in Section 4.4. The idea is to design a bank of local robust FDI schemes in a distributed manner with each LFDI scheme corresponding to a subsystem. Time-varying thresholds are selected by explicitly accounting for the effect of uncertainties and faults in shared interconnections that cannot be isolated in the corresponding subsystem. In this way, robustness of the LFDI schemes to false alarms is achieved. Also, in the case of faults in the shared interconnections, the distributed architecture of the FDI framework allows the other FDI schemes to function as intended. This is achieved via utilizing healthy estimation of faulty shared interconnection and introducing a new concept called detectability index that measures the probability of a residuals breaching its threshold when it is expected. The detectability and isoability conditions are rigorously derived for the distributed FDI scheme. Next, effectiveness of the proposed methodology is shown via application to a reactor-separator process subject to uncertainty and measurement noise in Section 4.5. Finally, some

concluding remarks are presented in Section 4.6.

4.2 Preliminaries

Consider a networked uncertain nonlinear system composed of M subsystems with the ith subsystem described by

$$\begin{aligned} \dot{x}_i &= f_i(x_i) + G_i(x_i)(u_i + u_{f,i}) + \theta_i(x_i, u_i, t) \\ &\quad + I_i(x_i, u_i, \tilde{y}_i) + d_i(x_i, u_i, \tilde{x}_i, \tilde{y}_i, t) \\ y_i &= h_i(x_i) + y_{f,i} \end{aligned} \tag{4.1}$$

where $x_i \in \mathcal{X}_i \subset \mathbb{R}^{n_i}$ denotes the vector of state variables corresponding to the ith subsystem, with $\mathcal{X}_i$ being a compact set of the admissible state values, $u_i = [u_1, \ldots, u_{m_i}]^{\mathrm{T}} \in \mathbb{R}^{m_i}$ denotes the vector of prescribed control inputs corresponding to the ith subsystem, taking values in a nonempty compact convex set $\mathcal{U}_i \subseteq \mathbb{R}^{m_i}$, $u_{f,i} = [u_{f_1}, \ldots, u_{f_{m_i}}]^{\mathrm{T}} \in \mathbb{R}^{m_i}$ denotes the unknown fault vector for the actuators corresponding to the ith subsystem, θ_i denotes the unstructured uncertainty (unmodeled dynamics) corresponding to the ith subsystem with $\|\theta_i(x_i, u_i, t)\| \leq \bar{\theta}_i$, where $\bar{\theta}_i$ is a known positive constant, $y_i = [y_1, \ldots, y_{p_i}]^{\mathrm{T}} \in \mathbb{R}^{p_i}$ denotes the vector of output variables corresponding to the ith subsystem, $y_{f,i} = [y_{f_1}, \ldots, y_{f_{p_i}}]^{\mathrm{T}} \in \mathbb{R}^{p_i}$ denotes the unknown fault vector for the sensors corresponding to the ith subsystem, $G_i(x_i) = [g_{i1}(x_i), \ldots, g_{im_i}(x_i)]$. Furthermore, the vector field $I_i(x_i, u_i, \tilde{y}_i)$ represents the known part of direct interconnection between the ith subsystem and the other subsystems, $d_i(x_i, u_i, \tilde{x}_i, \tilde{y}_i, t)$ represents the unknown part of direct interconnection between the ith subsystem and the other subsystems, with $\|d_i(x_i, u_i, \tilde{x}_i, \tilde{y}_i, t)\| \leq \bar{d}_i$,

where $\bar{d}_i$ is a known positive constant. Note that $\tilde{x}_i \in \mathbb{R}^{\tilde{n}_i}$ and $\tilde{y}_i = [\tilde{y}_1, \ldots, \tilde{y}_{\tilde{p}_i}] \in \mathbb{R}^{\tilde{p}_i}$ denote the state variables and the output variables of the the neighboring subsystems that affect the ith subsystem, respectively and $\tilde{y}_{f,i} = [\tilde{y}_{f_1}, \ldots, \tilde{y}_{f_{\tilde{p}_i}}]^{\mathrm{T}} \in \mathbb{R}^{\tilde{p}_i}$ denotes the unknown fault vector for the sensors corresponding to the output variables from the subsystems that affects the *ith* subsystem (these subsystems are henceforth denoted as 'neighboring' subsystems). In the rest of this manuscript, we refer to $\tilde{y}_i$ and $\tilde{y}_{f,i}$ as shared interconnections and faulty shared interconnection with the ith subsystem, respectively.

Due to the presence of physical constraints, the actual input $u_i + u_{f,i}$ implemented to the subsystem takes values from the set $\mathcal{U}_i$ as well. The inputs are implemented in a discrete fashion, with sampling time Δ_i. In this work, it is assumed that the sampling time for all of the subsystems is the same and as a result, the notation Δ is used as sampling time of all the subsystems.

Remark 4.1. Note that in this work (as with other existing results in the literature see e.g., Ferrari *et al.* (2012); Keliris *et al.* (2015)) the direct interconnection terms are not allowed to be a function of neighboring subsystems inputs u_j. This follows from the assumption that each subsystem can be operated with a local controller (possibly utilizing measurements from the other subsystems).

4.3 Boundedness of estimation error in the presence of uncertainty and exchange of information for networked systems

State estimates are required for the design of the proposed FDI scheme and as a result, boundedness of estimation error must first be established. In this section we use a high gain observer, and establish boundedness of the estimation error in the presence of uncertainty and exchange of information. To this end, we consider the ith subsystem described by Eq. 4.1 under fault free conditions, satisfying Assumptions 4.1-4.3:

Assumption 4.1. The functions $f_i : \mathbb{R}^{n_i} \to \mathbb{R}^{n_i}$, $g_{ik} : \mathbb{R}^{n_i} \to \mathbb{R}^{n_i}$, $k = 1, \ldots, m_i$, $\theta_i(x_i, u_i, t) : \mathbb{R}^{n_i} \to \mathbb{R}^{n_i}$, $I_i(x_i, u_i, \tilde{y}_i) : \mathbb{R}^{n_i} \to \mathbb{R}^{n_i}$, $d_i(x_i, u_i, \tilde{x}, \tilde{y}_i, t) : \mathbb{R}^{n_i} \to \mathbb{R}^{n_i}$ and $h_i : \mathbb{R}^{n_i} \to \mathbb{R}^{p_i}$ are smooth on their domains of definition, and $f_i(0) = 0$.

Assumption 4.2. For the ith subsystem of the network presented by Eq. 4.1, there exists a positive definite $\mathcal{C}^2$ function $V_i : \mathbb{R}^{n_i} \to \mathbb{R}$ such that for any $x_i \in \Omega_{c_i} := \{x_i \in \mathbb{R}^n : V_i(x_i) \leq c_i\}$, where c_i is a positive real number, the following inequality holds:

$$L_{f_i} V_i(x_i) + L_{g_i} V_i(x_i) u_i(x_i) + L_{\theta_i} V_i(x_i) + L_{I_i} V_i(x_i) + L_{d_i} V_i(x_i) \leq -\alpha_i(V_i(x_i)) \quad (4.2)$$

where $L_{g_i} V_i(x) = [L_{g_{i1}} V_i(x_i), \ldots, L_{g_{im}} V_i(x_i)]$, $u_i : \Omega_{c_i} \to \mathcal{U}_i$ is a state feedback control law and α_i is a class $\mathcal{K}$ function.

Remark 4.2. Note that there is no general procedure for construction of robust control Lyapunov functions (RCLFs) for nonlinear uncertain systems of the form of

Eq. 4.1. However, there are several class of systems for which such procedure exists (see e.g., Freeman and Kokotovic (2008)). Also, recent results demonstrate the ability to verify CLFs as RCLFs (see e.g., Mahmood *et al.* (2008a)). More importantly, note that Assumption 4.2 simply states that a control design is in place to handle the uncertainly in the system, and does not require the knowledge of the specific Lyapunov function for the FDI design.

Assumption 4.3. There exist integers ω_{ik}, $k = 1, \ldots, p_i$, with $\sum_{i=1}^{p_i} \omega_{ik} = n_i$, and a coordinate transformation $\zeta_i = T_i(x_i, u_i, t) = T_i'(x_i, u_i) + T_{\theta_i}(x_i, u_i, t)$ such that if $u_i = \bar{u}_i$, where $\bar{u}_i \in \mathcal{U}_i$ is a constant vector, then the representation of the system of Eq. 4.1 in the ζ_i coordinate takes the following form:

$$\begin{aligned} \dot{\zeta}_i &= A_i\zeta_i + B_i\phi_i(\zeta_i, \bar{u}_i) + \eta_i(\zeta_i, \bar{u}_i, t) \\ y_i &= C_i\zeta_i \end{aligned} \tag{4.3}$$

where $\zeta_i = [\zeta_{i1}, \ldots, \zeta_{ip_i}]^{\mathrm{T}} \in \mathbb{R}^{n_i}$, $A_i = \text{blockdiag}[A_{i1}, \ldots, A_{ip_i}]$, $B_i = \text{blockdiag}[B_{i1}, \ldots, B_{ip_i}]$, $C_i = \text{blockdiag}[C_{i1}, \ldots, C_{ip_i}]$, $T_i' = \phi_i = [\phi_{i1}, \ldots, \phi_{ip_i}]^{\mathrm{T}}$, $T_{\theta_i} = \eta_i = [\eta_{i1}, \ldots, \eta_{ip_i}]^{\mathrm{T}}$, $\zeta_{ik} = [\zeta_{ik,1}, \ldots, \zeta_{ik,\omega_{ik}}]^{\mathrm{T}}$, $A_{ik} = \begin{bmatrix} 0 & I_{\omega_{ik}-1} \\ 0 & 0 \end{bmatrix}$, with $I_{\omega_{ik}-1}$ being a $(\omega_{ik}-1) \times (\omega_{ik}-1)$ identity matrix, $B_{ik} = [0_{\omega_{ik}-1}^{\mathrm{T}}, 1]^{\mathrm{T}}$, with $0_{\omega_{ik}-1}$ being a vector of zeros of dimension $\omega_{ik} - 1$, $C_{ik} = [1, 0_{\omega_{ik}-1}^{\mathrm{T}}]$, $\phi_{ik}(\zeta, \bar{u}) = \phi_{ik,\omega_{ik}}(\zeta, \bar{u})$, with $\phi_{ik,\omega_{ik}}(\zeta_i, \bar{u}_i)$ defined through the successive differentiation of $h_i(x_i)$: $\phi_{ik,1}(\zeta_i, \bar{u}_i) = h_i(x_i)$ and $\phi_{ik,j}(\zeta_i, \bar{u}_i) = \frac{\partial \phi_{ik,j-1}}{\partial x_i}[f_i(x_i) + g_{ik}(x_i)\bar{u}_i + I_i(x_i, \bar{u}_i, \tilde{y}_i)]$ and $\eta_{ik}(\zeta_i, \bar{u}_i, t) = \eta_{ik,\omega_{ik}}(\zeta_i, \bar{u}_i, t)$, with $\eta_{ik,\omega_{ik}}(\zeta_i, \bar{u}_i, t)$ defined: $\eta_{ik,1}(\zeta_i, \bar{u}_i, t) = 0$ and $\eta_{ik,j}(\zeta_i, \bar{u}_i, t) = \frac{\partial \phi_{ik,j-1}}{\partial x}[\theta_i(x_i, u_i, t) + d_i(x_i, u_i, \tilde{x}_i, \tilde{y}_i, t)]$. Furthermore, $T_i' : \mathbb{R}^{n_i} \times \mathcal{U}_i \to \mathbb{R}_i^n$, $T_{\theta_i} : \mathbb{R}^{n_i} \times \mathcal{U}_i \to \mathbb{R}_i^n$, $T_i'^{-1} : \mathbb{R}^{n_i} \times \mathcal{U}_i \to \mathbb{R}^{n_i}$, and $T_{\theta_i}^{-1} : \mathbb{R}^{n_i} \times \mathcal{U}_i \to \mathbb{R}^{n_i}$ are $\mathcal{C}^1$ functions on

their domains of definition, $\eta_{i,j}$ denotes the uncertainty in the new coordinate system and $\|\eta_{i,j}(\zeta_i, u_i, t)\| \leq \bar{\eta}_{i,j}$, where $\bar{\eta}_{i,j}$ is a known positive constant.

Remark 4.3. Note the above assumptions simply state that the assumptions typically utilized for centralized robust FDI design in Chapter 3 continue to hold in the networked system. Furthermore, these assumptions can be verified off-line either in a centralized or a decentralized fashion.

We next describe the utilized high-gain observer formulation, which is coupled with sample-and-hold control. In the closed-loop system, the input is prescribed at discrete times $t_k = k\Delta$, $k = 0, \ldots, \infty$, with Δ being the hold-time of the control action. For $t \in [t_k, t_{k+1})$, the observer is formulated as follows (see Chapter 3 for more on this):

$$\begin{aligned} \dot{\hat{\zeta}}_i &= A_i\hat{\zeta}_i + H_i(y_i - C_i\hat{\zeta}_i) \\ \hat{\zeta}_i(t_k) &= T_i'(\hat{x}_i(t_k), u_i(t_k)) \end{aligned} \tag{4.4}$$

where $\hat{x}_i$ and $\hat{\zeta}_i$ denote the estimates of x_i and ζ_i, respectively, $H_i = \text{blockdiag}[H_{i1}, \ldots, H_{ip_i}]$ is the observer gain, $H_{ik} = \left[\frac{a_{ik,1}}{\varepsilon}, \ldots, \frac{a_{ik,\omega_{ik}}}{\varepsilon^{\omega_{ik}}}\right]^{\mathrm{T}}$, with $s^{\omega_{ik}} + a_{ik,1}s^{\omega_{ik}-1} + \cdots + a_{ik,\omega_{ik}} = 0$ being a Hurwitz polynomial and ε_i being a positive constant to be specified, $\hat{x}_i(t_k) = T_i'^{-1}(\hat{\zeta}_i(t_k^-), u_i(t_{k-1})))$ for $k = 1, \ldots, \infty$. Note that Eq. 4.4 results from choosing $\phi_{0_i}(\cdot,\cdot)$, the nominal model of ϕ_i, as zero. This choice satisfies the global boundedness requirement (see Du and Mhaskar (2014) and Shahnazari and Mhaskar (2016)). The initial state of the observer is denoted by $\hat{x}_{0_i} := \hat{x}_i(0)$, which takes values from any compact set $\mathcal{Q}_i \subseteq \mathbb{R}^{n_i}$. In the transformed coordinate, the state estimate $\hat{\zeta}_i$ is re-initialized at the discrete times to account for the possible discrete changes in the input and ensuring that the resulting state estimates remain continuous.

Preparatory to the presentation of results on the convergence of the observer, we first state an important property of the scaled estimation error. To this end, let $D_i = \text{blockdiag}[D_{i1}, \ldots, D_{ip_i}]$, where $D_{ik} = \text{diag}[\varepsilon^{\omega_{ik}-1}, \ldots, 1]$, and define the scaled estimation error $e_i = D_i^{-1}(\zeta_i - \hat{\zeta}_i) \in \mathbb{R}^{n_i}$. For $t \in [t_k, t_{k+1})$, the scaled estimation error evolves as follows:

$$\begin{aligned} \varepsilon_i \dot{e}_i &= A_{0_i} e_i + \varepsilon_i B_i \phi_i(\zeta_i, u_i(t_k)) + \varepsilon_i \eta_i(\zeta_i, u_i(t_k), t) \\ e_i(t_k) &= D_i^{-1}[T_i(x_i(t_k), u_i(t_k), t) - T_i'(\hat{x}_i(t_k), u_i(t_k), t)] \end{aligned} \tag{4.5}$$

where $A_{0_i} = \text{blockdiag}[A_{0,i1}, \ldots, A_{0,ip_i}]$, $A_{0,ik} = [a_{ik}, b_{ik}]$, $a_{ik} = [-a_{ik,1}, \ldots, -a_{ik,\omega_{ik}}]^{\mathrm{T}}$, and $b_{ik} = [I_{\omega_{ik}-1}, 0_{\omega_{ik}-1}]^{\mathrm{T}}$.

Applying the change of time variable $\tau_i = \frac{t}{\varepsilon_i}$ and setting $\varepsilon_i = 0$, the boundary-layer system is given by

$$\frac{de_i}{d\tau_i} = A_{0_i} e_i \tag{4.6}$$

For the boundary-layer system, we define a Lyapunov function $W(e_i) = e_i^{\mathrm{T}} P_{0_i} e_i$, where P_{0_i} is the symmetric positive definite solution of the Lyapunov function $A_{0_i}^{\mathrm{T}} P_{0_i} + P_{0_i} A_{0_i} = -I_i$. Let $\lambda_{\min_i}$ and $\lambda_{\max_i}$ denote the minimum and maximum eigenvalues of P_{0_i}, respectively. Proposition 4.1 below is similar to a result obtained in the literature (see e.g., Atassi and Khalil (1999); Du and Mhaskar (2014)), and hence stated without proof.

Proposition 4.1. Consider the ith subsystem of the network presented by Eq. 4.1, for which Assumptions 4.1 and 4.3 hold. If $x_{0_i} := x_i(0) \in \Omega_{b_i}$, where $0 < b_i < c_i$, then given $b_i' \in (b_i, c_i)$, there exists a finite time t_{e_i}, independent of ε_i, such that $x_i(t) \in \Omega_{b_i'}$ for all $t \in [0, t_{e_i}]$. Furthermore, there exists $\sigma_i > 0$, independent of ε_i, such that for any $e_i(t) \in \mathcal{W}_{o_i} := \{e_i \in \mathbb{R}^{n_i} : W_i(e_i) \geq \sigma_i \varepsilon_i^2\}$ and $x_i(t) \in \Omega_{c_i}$, $\dot{W}_i \leq -\frac{1}{2\varepsilon_i}\|e_i\|^2$.

Theorem 4.1 formalizes the convergence property of observer design and stability of the ith closed loop subsystem in the presence of uncertainty. The results below is similar to the result for a centralized design presented in Chapter 3, and hence is stated without proof.

Theorem 4.1. *Consider the ith subsystem of the network presented by Eq. 4.1, for which Assumptions 4.1-4.3 hold, under a stabilizing local control law u_i. Given any $0 < b_i < c_i$, $d_i > 0$, $d_i' > 0$ and $\bar{\theta}_i$, there exist $\Delta_i^*(\bar{\theta}_i) > 0$ and $\varepsilon_i^*(\bar{\theta}_i) > 0$ such that if $\Delta \in (0, \Delta_i^*(\bar{\theta}_i)]$, $\varepsilon_i \in (0, \varepsilon_i^*(\bar{\theta}_i)]$, and $x_{0_i} \in \Omega_{b_i}$, then 1) there exists an integer $k_i' > 0$ such that $\|\hat{x}_i(t_{k_i}) - x_i(t_{k_i})\| \leq d_i' \ \forall \ t_{k_i} \geq t_{k_i'}$, and 2) $x_i(t) \in \Omega_{c_i} \ \forall \ t_i \geq 0$ and $\limsup_{t\to\infty} \|x_i(t)\| \leq d_i$.*

4.4 Distributed fault detection and isolation design

The distributed FDI architecture is composed of M communicating LFDI schemes with each monitoring one of the network subsystems. The communications are limited to the LFDI schemes corresponding to the neighboring subsystems. The ith LFDI scheme requires the prescribed value of inputs, outputs measurements of the ith subsystem and measurements of only the outputs from the neighbor subsystems that affect the ith subsystem. Note that the measurements of the outputs from the neighbor subsystems that affect the ith subsystem are communicated by local information transmitter in each subsystem that can be a part of the LFDI schemes if a model based controller is not in place. The LFDI schemes also communicate to share

the available information regarding the diagnosed faults in the shared interconnections with the ith subsystem. Faults are classified into two categories based on the fault location:

1. Local faults: This type of faults does not take place in the shared sensors with other subsystems.

2. Distributed faults: This type of faults takes place in the shared sensors with other subsystems.

The local faults are diagnosed using the corresponding LFDI scheme, if they can be isolated by the corresponding LFDI scheme, i.e. they are distinguishable and isolable in their corresponding LFDI scheme (for definition of distinguishable fault scenarios and isolable faults see e.g., Shahnazari *et al.* (2016); Zhang *et al.* (2010b)). In this work, it is assumed that all of the possible fault scenarios are distinguishable. The distributed faults are diagnosed using the LFDI schemes corresponding to their subsystem of origin. The communication between LFDI schemes allows the LFDI schemes to function as intended, in the presence of distributed faults.

4.4.1 Residual generation

We consider the case where at most two faults take place in the ith subsystem. By using the principles of combinatorics, the number of possible distinguishable fault scenarios corresponding to the ith subsystem is $n_{f_i} = m_i + \frac{m_i(m_i-1)}{2} + p_i + \frac{p_i(p_i-1)}{2} + m_i p_i$, where, for the ith subsystem, m_i and p_i denote the number of actuator and sensor faults originated in the ith subsystem, respectively. Residuals are designed using the

methodology presented in Du *et al.* (2013) for centralized design, albeit for the local subsystem.

Before proceeding to present the FDI mechanism, we need to employ the following assumption:

Assumption 4.4. The ith subsystem state vector x_i remains bounded before and after fault occurrence i.e., there exist a positive constant d_{g_i} such that $\|x_i\| \leq d_{g_i}$, $\forall t > 0$

Remark 4.4. Note that Assumption 4.4 states that the proposed FDI methodology remains applicable under any stabilizing output feedback controller implemented in a discrete fashion as long as each subsystem state evolves within a compact set before and after fault occurrence.

In particular, the residual for a fault scenario is defined as the norm of the difference between the state prediction and the state estimate for the component of the ith subsystem that does not require the value of the corresponding actuator/sensor in the calculations (see e.g., Du *et al.* (2013)). A residual dedicated to a particular fault scenario is designed so it is only insensitive to that particular fault scenario but sensitive to the other fault scenarios (we thus design so-called generalized residuals). To this end, let Θ_{f,j_i} denote the fault vector (sensor/and or actuator) for the jth fault scenario where $j = 1, \ldots, n_{f_i}$, and $\bar{\Theta}_{f,j_i}$ the remaining fault variable vector (the remaining u_f and y_f variables). Specifically, let u_{f,j_i} and y_{f,j_i} denote the vectors of input and output variables subject to faults Θ_{f,j_i}, respectively. Let $\bar{u}_{f,j_i}$ and $\bar{y}_{f,j_i}$ denote the vectors of the rest of the input and output variables, respectively. To design the insensitive residual to Θ_{f,j_i}, there must exist at least one observer that can be

designed without using the inputs subject to fault or the faulty measurements that belong to Θ_{f,j_i}. Assumption 4.5 establishes this.

Assumption 4.5. Du *et al.* (2013) Assumption 4.3 holds for the ith subsystem of networked system described Eq. 4.1, with $\bar{u}_{f,j_i}$ and $\bar{y}_{f,j_i}$ being the vectors of input and output variables, respectively, $j = 1, \ldots, n_{f_i}$ where n_{f_i} is the number of possible fault scenarios corresponding to the ith subsystem.

Under Assumption 4.5, the state observer for the jth fault scenario is designed as follow :

$$\begin{aligned} \dot{\hat{\zeta}}^{l,i} &= A^l \hat{\zeta}^{l,i} + H^{l,i}(\bar{y}_{f,j_i} - C^{l,i}\hat{\zeta}^j) \\ \dot{\hat{\zeta}}^{l,i}(t_k) &= T'^{l,i}(\hat{x}^j(t_k), \bar{u}_{f,j_i}(t_k)) \end{aligned} \tag{4.7}$$

where l represents the lth observer designed for the ith FDI scheme with $l = 1, \ldots, p_i + \frac{p_i(p_i-1)}{2}$.

To define residuals, we need to calculate the expected plant trajectories. To this end, we consider a component of ith subsystem of Eq. 4.1 for which the state variables are all of those such that no inputs in $u_{f_{j_i}}$ appear on the right-hand side of the corresponding ODE's. Let $x_{sub,j,i}$ denote the vector of state variables for the component, and $\bar{x}_{sub,i,i}$ the vector of the rest of the state variables. Without loss of generality, the model of the component can be described as follows:

$$\begin{aligned} \dot{x}_{sub,j,i} = {} & f_{sub,j,i}([x^{\mathrm{T}}_{sub,j,i}, \bar{x}^{\mathrm{T}}_{sub,j,i}]^{\mathrm{T}}) + G_{sub,j,i}([x^{\mathrm{T}}_{sub,j,i}, \bar{x}^{\mathrm{T}}_{sub,j}]^{\mathrm{T}})\bar{u}_{f,j,i} \\ & + I_{sub,j,i}([x^{\mathrm{T}}_{sub,j,i}, \bar{x}^{\mathrm{T}}_{sub,j,i}]^{\mathrm{T}}, \bar{u}_{f,j,i}, \tilde{y}_{sub,j,i}) + \theta_{sub,j,i}([x^{\mathrm{T}}_{sub,j,i}, \bar{x}^{\mathrm{T}}_{sub,j,i}]^{\mathrm{T}}, \bar{u}_{f,j,i}, t) \\ & + d_{sub,j,i}([x^{\mathrm{T}}_{sub,j,i}, \bar{x}^{\mathrm{T}}_{sub,j,i}]^{\mathrm{T}}, \bar{u}_{f,j,i}, \tilde{x}_{sub,j,i}, \tilde{y}_{sub,j,i}, t) \end{aligned} \tag{4.8}$$

where $f_{sub,j,i}(\cdot)$, $G_{sub,j,i}(\cdot)$, $I_{sub,j,i}(\cdot)$, $\theta_{sub,j,i}(\cdot)$ and $d_{sub,j,i}(\cdot)$ are appropriately defined. For each faulty scenario, the expected component trajectory is computed using the known part of the system model and the state estimates generated by the lth observer that does not require values of the variables included in the fault vector $\Theta_{f,j,i}$. Specifically, for $t \in [t_{k-T}, t_k)$, a prediction model is designed as follows:

$$\begin{aligned}\dot{\tilde{x}}_{sub,j,l,i} = f_{sub,j,i}([\tilde{x}^{\mathrm{T}}_{sub,j,l,i}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,j,l,i}]^{\mathrm{T}}) + G_{sub,j,i}([\tilde{x}^{\mathrm{T}}_{sub,j,l,i}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,j,l,i}]^{\mathrm{T}})\bar{u}_{f,j,i} \\ + I_{sub,j,i}([[\tilde{x}^{\mathrm{T}}_{sub,j,l,i}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,j,l,i}]^{\mathrm{T}}, \bar{u}_{f,j,i}, \tilde{y}_{sub,j,i})\end{aligned} \tag{4.9}$$

where $\tilde{x}_{sub,j,l,i}$ is the state of the prediction model, $\hat{\tilde{x}}_{sub,j,l,i}$ is the estimate of $\bar{x}_{sub,j,i}$ provided by the lth observer, and T is the prediction horizon: $T = 1$ if $0 < t_k \leq t_{k'}$; $T = k - k'$ if $t_{k'} < t_k \leq t_{k'+T_p}$; and $T = T_p$ if $t_k > t_{k'+T_p}$, with a positive integer T_p being a chosen prediction horizon. The initial condition for the prediction model is the state estimate at time t_{k-T}: $\tilde{x}_{sub,j,l,i}(t_{k-T}) = \hat{x}_{sub,j,l,i}(t_{k-T})$. Let $\tilde{x}_{sub,j,l,i}(t_k)$ denote the prediction for the state vector $x_{sub,j,i}$ at time t_k. By solving Eq. 4.9, the state prediction at time t_k is obtained.

For the jth faulty scenario in the ith subsystem, the residual (at the time instance t_{k+1}) is defined as follows:

$$r_{j,k+1,i} = \|\tilde{x}_{sub,j,l,i}(t_{k+1}) - \hat{x}_{sub,j,l,i}(t_{k+1})\| \tag{4.10}$$

Note that in this work we assume all of the fault scenarios are distinguishable. Assumption 4.6 presents this. To this end, let $\Theta_{j,i}$ and $r_{j,i,ins}$ denote the jth fault scenario corresponding to the ith subsystem and the set of insensitive residuals corresponding to the $\Theta_{j,i}$, respectively.

Assumption 4.6. The conditions presented for centralized FDI Shahnazari *et al.* (2016) hold for all of the fault scenarios corresponding to the ith subsystem i.e., $\Theta_{j,i}$ is distinguishable if and only if there exists a one-to-one mapping between every fault scenario and $r_{j,i,ins}$, where $j, \in \{1, \ldots, n_{f,i}\}$.

Remark 4.5. Note that Assumption 4.6 states there exist enough analytical redundancy in the system model that enables isolation of all possible fault scenarios in the absence of uncertainty regardless of the type of estimation scheme being used for generating residuals. However, in the presence of uncertainty, satisfying Assumption 4.6 is only a necessary condition for fault isolation and a fault can be diagnosed only if its effect does not get compensated by uncertainty and as a result, all of the corresponding sensitive residuals breach their thresholds. This is due to the inherent trade-off between robustness and uncertainty that is discussed later in the manuscript (see the existing results in the literature for more discussion on this issue (see e.g., Chapter 3, Zhang *et al.* (2010b); Zhang (2011))).

4.4.2 Threshold design

In this section, we present the threshold design. For each LFDI scheme, thresholds are selected by explicitly accounting for the effect of uncertainty and updated upon isolation of a distributed fault affecting the corresponding LFDI scheme. In this way, robustness of the LFDI scheme is guaranteed with respect to uncertainty and distributed faults affecting the corresponding subsystem. To this end, consider the scenario where a fault in one of the sensors corresponding to the $\tilde{y}_i$ is isolated by its corresponding LFDI scheme and let us assume that the origin of the fault in $\tilde{y}_i$ is the qth subsystem. In this case, the qth LFDI scheme notifies the ith LFDI

scheme and sends the healthy estimates of $\tilde{y}_i$ to the ith LFDI scheme to replace the faulty measurements. Then thresholds of the ith LFDI scheme are updated using the provided healthy estimates to make sure the ith LFDI scheme functions as intended. Thus the detectability and isolability properties of the ith LFDI scheme are retained in the presence of distributed faults.

Before defining thresholds corresponding to each residual, we need the following result presented as Lemma 4.1. To this end, let $t_{f_k,dis,i}$ denote the time of isolation of distributed fault $\Theta_{f,dis,i}$ in $\tilde{y}_{i,\Theta_{f,dis,i}}$ by the corresponding LFDI scheme, where $\tilde{y}_{\Theta_{f,dis,i},i}$ denotes the vector of shared outputs affecting the ith subsystem subject to fault $\Theta_{f,dis,i}$, $\bar{\tilde{y}}_{\Theta_{f,dis,i},i}$ denotes the shared outputs affecting the ith subsystem that are not subject to to fault $\Theta_{f,dis,i}$ and $\hat{\tilde{y}}_{\Theta_{f,dis,i},i}$ denotes the healthy estimate of $\tilde{y}_{\Theta_{f,dis,i},i}$ provided by the corresponding LFDI scheme.

Lemma 4.1. Consider the ith subsystem of network presented by Eq. 4.1, for which Assumptions 4.1-4.6 hold and residuals for each sample time are defined as in Eq. 4.10. Then under fault free condition, for $t_{k'} \leq t_k \leq t_{f_k,dis,i}$ (see Theorem 4.1 for definition of k'):

$$r_{j,l,i,k+1} \leq L_{2,sub,j,l,i}\beta_3^{j,l,i}(\varepsilon)E_{s,sub,j,l,i} + \bar{\theta}_{sub,j,l,i}\Delta + E_p^{j,l,i} \tag{4.11}$$

and for $t_{f_k,dis} < t_k$

$$r_{j,l,i,k+1} \leq L_{2,sub,j,l,i}\beta_3^{j,l,i}(\varepsilon)E_{s,sub,j,l,i} + \bar{\theta}_{sub,j,l,i}\Delta + \hat{E}_p^{j,l,i} \tag{4.12}$$

where $E_{s,sub,j,l,i} = e^{-\alpha_{sub,j,l,i}\Delta}e_b^{*,j,i} + K^j_{B,sub,j,l,i}K^j_{\phi,sub,j,l,i}\frac{1-e^{-\alpha_{sub,j,l,i}\Delta}}{\alpha_{sub,j,l,i}} + \frac{1-e^{-\alpha_{sub,j,l,i}\Delta}}{\alpha_{sub,j,l,i}}$

$\bar{\eta}_{sub,j,l,i}$, *sub* and l refer to the corresponding subsystem and observer used for defining $r_{j,i}$, respectively, where j refers to jth residual, $e_b^{*,j}$ is the upper bound on $\|e^j_{sub,j,l,i}(t_k)\|$, $K_{B_{j,i}}$ is spectrum bound of the matrix $\|B^{j,i}\|$, $\alpha_{j,i}$ is the spectrum bound of the matrix $\frac{A_0^{j,i}}{\varepsilon}$ and $E_p^{j,i} = (M^{j,i} + M'^{j,i})\Delta + K_{\tilde{e}}^{j,i}$, where
$\|(\tilde{f}_{sub,j,i} - f_{sub,j,i}([x^{\mathrm{T}}_{sub,j,i}, \bar{x}^{\mathrm{T}}_{sub,j,i}]^{\mathrm{T}}))\| + \|(\tilde{G}_{sub,j,i} - G_{sub,j,i}([x^{\mathrm{T}}_{sub,j,i}, \bar{x}^{\mathrm{T}}_{sub,j,i}]^{\mathrm{T}}))\bar{u}_{f,j,i}\| + \bar{\theta}_{sub,j,i} \leq M^{j,i}$, $\|\tilde{I}_{sub,j,i} - I_{sub,j,i}([[x^{\mathrm{T}}_{sub,j,l,i}, \bar{x}^{\mathrm{T}}_{sub,j,l,i}]^{\mathrm{T}}, \bar{u}_{f,j,i}, \tilde{y}_{sub,j,i})\| + \bar{d}_{sub,j,i} \leq M'^{j,i}$, where $\tilde{f}_{sub,j,i} = f_{sub,j,i}([\tilde{x}^{\mathrm{T}}_{sub,j,l,i}, \hat{\bar{x}}^{\mathrm{T}}_{sub,j,l,i}]^{\mathrm{T}})$, $\tilde{G}_{sub,i} = G_{sub,i}([\tilde{x}^{\mathrm{T}}_{sub,i,j}, \hat{\bar{x}}^{\mathrm{T}}_{sub,i,j}]^{\mathrm{T}})$,
$\tilde{I}_{sub,i} = I_{sub,i}([[\tilde{x}^{\mathrm{T}}_{sub,j,l,i}, \hat{\bar{x}}^{\mathrm{T}}_{sub,j,l,i}]^{\mathrm{T}}, \bar{u}_{f,j,i}, \tilde{y}_{sub,j,i})$, $K_{\tilde{e}_{j,i}}$ is the upper bound on $\|\tilde{e}_{j,i,k}(t_k)\|$ which is constant, where $\tilde{e}_{j,i,k} = \tilde{x}_{sub,j,l,i} - x_{sub,j}$ and $\tilde{x}_{sub,j,l,i}$ is state of prediction model defined in Du *et al.* (2013), $\hat{E}_p^{j,i} = (M^{j,i} + \hat{M}'^{j,i})\Delta + \hat{K}_{\tilde{e}}^{j,i}$, $\|\tilde{I}_{sub,j,i} - I_{sub,j,i}([[x^{\mathrm{T}}_{sub,j,l,i}, \bar{x}^{\mathrm{T}}_{sub,j,l,i}]^{\mathrm{T}}, \bar{u}_{f,j,i}, \tilde{y}_{sub,j,i})\| + \bar{d}_{sub,j,i} \leq \hat{M}'^{j,i}$ where
$\tilde{I}_{sub,j,l,i} = I_{sub,j,l,i}([[\tilde{x}^{\mathrm{T}}_{sub,j,l,i}, \hat{\bar{x}}^{\mathrm{T}}_{sub,j,l,i}]^{\mathrm{T}}, \bar{u}_{f,j,i}, \hat{\bar{y}}_{\Theta_{f,dis,i},i}, \bar{\tilde{y}}_{\Theta_{f,dis,i},i}$ and $\|\tilde{e}_{j,i,k}(t_k)\| \leq \hat{K}_{\tilde{e}_{j,i}}$.

Proof. By using triangular inequality, Eq. 4.10 turns to the following form:

$$\begin{aligned} r_{i,k+1} &= \|\tilde{x}_{sub,i,j}(t_{k+1}) - \hat{x}_{sub,i,j}(t_{k+1})\| \\ &\leq \|\tilde{x}_{sub,i,j}(t_{k+1}) - x_{sub,i}(t_{k+1})\| + \|x_{sub,i}(t_{k+1}) - \hat{x}_{sub,i,j}(t_{k+1})\| \\ &\leq \sup\|\tilde{x}_{sub,i,j}(t_{k+1}) - x_{sub,i}(t_{k+1})\| + \sup\|x_{sub,i}(t_{k+1}) - \hat{x}_{sub,i,j}(t_{k+1})\| \end{aligned} \tag{4.13}$$

Consider the prediction model corresponding to fault Θ_{f_i}:

$$\begin{aligned} \dot{\tilde{x}}_{sub,j,l,i} &= f_{sub,j,i}([\tilde{x}^{\mathrm{T}}_{sub,j,l,i}, \hat{\bar{x}}^{\mathrm{T}}_{sub,j,l,i}]^{\mathrm{T}}) + G_{sub,j,i}([\tilde{x}^{\mathrm{T}}_{sub,j,l,i}, \hat{\bar{x}}^{\mathrm{T}}_{sub,j,l,i}]^{\mathrm{T}})\bar{u}_{f,j,i} \\ &\quad + I_{sub,j,i}([[\tilde{x}^{\mathrm{T}}_{sub,j,l,i}, \hat{\bar{x}}^{\mathrm{T}}_{sub,j,l,i}]^{\mathrm{T}}, \bar{u}_{f,j,i}, \tilde{y}_{sub,j,dis,dis,i}, \tilde{y}_{sub,j,dis,indis,i}) \\ &\quad + d_{sub,j,i}([x^{\mathrm{T}}_{sub,j,i}, \bar{x}^{\mathrm{T}}_{sub,j,i}]^{\mathrm{T}}, \bar{u}_{f,j,i}, \tilde{x}_{sub,j,i}, \tilde{y}_{sub,j,i}, t) \end{aligned} \tag{4.14}$$

where $\tilde{x}_{sub,i,j}$ is the state of the prediction model, $\hat{\bar{x}}_{sub,i,j}$ is the estimate of $\bar{x}_{sub,i}$

provided by the corresponding observer.

By defining $\tilde{e}_{i,k} = \tilde{x}_{sub,i,j} - x_{sub,i}$, for $t_{k'} \leq t_k \leq t_{f_k,dis}$, we obtain:

$$\begin{aligned}\dot{\tilde{e}}_{i,k} &= f_{sub,i}([\tilde{x}^{\mathrm{T}}_{sub,i,j}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,i,j}]^{\mathrm{T}}) - f_{sub,i}([x^{\mathrm{T}}_{sub,i}, \bar{x}^{\mathrm{T}}_{sub,i}]^{\mathrm{T}}) + (G_{sub,i}([\tilde{x}^{\mathrm{T}}_{sub,i,j}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,i,j}]^{\mathrm{T}}) \\ &\quad - G_{sub,i}([x^{\mathrm{T}}_{sub,i}, \bar{x}^{\mathrm{T}}_{sub,i}]^{\mathrm{T}}))\bar{u}_{f,i} - \theta_{sub,i}([x^{\mathrm{T}}_{sub,i}, \bar{x}^{\mathrm{T}}_{sub,i}]^{\mathrm{T}}, \bar{u}_{f,i}, t) \\ &\quad + I_{sub,j,i}([[\tilde{x}^{\mathrm{T}}_{sub,j,l,i}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,j,l,i}]^{\mathrm{T}}, \bar{u}_{f,j,i}, \tilde{y}_{sub,j,i}) \\ &\quad - I_{sub,j,i}([[x^{\mathrm{T}}_{sub,j,l,i}, \bar{x}^{\mathrm{T}}_{sub,j,l,i}]^{\mathrm{T}}, \bar{u}_{f,j,i}, \tilde{y}_{sub,j,i}) \\ &\quad - d_{sub,j,i}([x^{\mathrm{T}}_{sub,j,i}, \bar{x}^{\mathrm{T}}_{sub,j,i}]^{\mathrm{T}}, \bar{u}_{f,j,i}, \tilde{x}_{sub,j,i}, \tilde{y}_{sub,j,i}, t)\end{aligned} \tag{4.15}$$

For sake of brevity, we define $\tilde{f}_{sub,j,l,i} = f_{sub,j,l,i}([\tilde{x}^{\mathrm{T}}_{sub,j,l,i}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,j,l,i}]^{\mathrm{T}})$, $\tilde{G}_{sub,j,l,i} = G_{sub,j,l,i}([\tilde{x}^{\mathrm{T}}_{sub,j,l,i}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,j,l,i}]^{\mathrm{T}})$ and $\tilde{I}_{sub,j,l,i}$ $= I_{sub,j,i}([[\tilde{x}^{\mathrm{T}}_{sub,j,l,i}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,j,l,i}]^{\mathrm{T}}, \bar{u}_{f,j,i}, \tilde{y}_{sub,j,i})$. By integration, we get:

$$\begin{aligned}\tilde{e}_{j,i,k}(t_{k+1}) &= \int_{t_k}^{t_{k+1}} (\tilde{f}_{sub,j,i} - f_{sub,j,i}([x^{\mathrm{T}}_{sub,i}, \bar{x}^{\mathrm{T}}_{sub,j,i}]^{\mathrm{T}}))d\tau \\ &\quad + \int_{t_k}^{t_{k+1}} (\tilde{G}_{sub,j,i} - G_{sub,j,i}([x^{\mathrm{T}}_{sub,j,i}, \bar{x}^{\mathrm{T}}_{sub,j,i}]^{\mathrm{T}}))\bar{u}_{f,j,i} - \theta_{sub,j,i} + \tilde{I}_{sub,j,i} \\ &\quad - I_{sub,j,i}([[x^{\mathrm{T}}_{sub,j,l,i}, \bar{x}^{\mathrm{T}}_{sub,j,l,i}]^{\mathrm{T}}, \bar{u}_{f,j,i}, \tilde{y}_{sub,j,i}) \\ &\quad - d_{sub,j,i}([x^{\mathrm{T}}_{sub,j,i}, \bar{x}^{\mathrm{T}}_{sub,j,i}]^{\mathrm{T}}, \bar{u}_{f,j,i}, \tilde{x}_{sub,j,i}, \tilde{y}_{sub,j,i}, t)d\tau + \tilde{e}_{j,i,k}(t_k)\end{aligned} \tag{4.16}$$

Under Assumptions 4.1 and 4.3, and by applying the triangular inequality, we obtain:

$$\|\tilde{e}_{j,i,k}(t_{k+1})\| \leq (M^{j,i} + M'^{j,i})\Delta + K_{\tilde{e}_{j,i}} \tag{4.17}$$

where $\|(\tilde{f}_{sub,j,i} - f_{sub,j,i}([x^{\mathrm{T}}_{sub,j,i}, \bar{x}^{\mathrm{T}}_{sub,j,i}]^{\mathrm{T}}))\| + \|(\tilde{G}_{sub,j,i} - G_{sub,j,i}([x^{\mathrm{T}}_{sub,j,i}, \bar{x}^{\mathrm{T}}_{sub,j,i}]^{\mathrm{T}}))\bar{u}_{f,j,i}\| + \bar{\theta}_{sub,j,i} \leq M^{j,i}$, $\|\tilde{I}_{sub,j,i} - I_{sub,j,i}([[x^{\mathrm{T}}_{sub,j,l,i}, \bar{x}^{\mathrm{T}}_{sub,j,l,i}]^{\mathrm{T}}, \bar{u}_{f,j,i}, \tilde{y}_{sub,j,i})\| + \bar{d}_{sub,j,i} \leq M'^{j,i}$ and $\|\tilde{e}_{j,i,k}(t_k)\| \leq K_{\tilde{e}_{j,i}}$. For $t_k > t_{f_k,dis}$, by replacing part of $\tilde{y}_{sub,j,i}$ that is subject to

distributed fault $\Theta_{f,dis,i}$ in $\tilde{I}_{sub,j,i}$ of Eq. 4.15, integration from $t_{k,dis}$ to $t_{k,dis}+1$ and using triangular inequality, we get:

$$\|\tilde{e}_{j,i,k}(t_{k+1})\| \leq (M^{j,i} + \hat{M}'^{j,i})\Delta + \hat{K}_{\tilde{e}_{j,i}} \tag{4.18}$$

where $\tilde{I}_{sub,j,l,i} = I_{sub,j,l,i}([\tilde{x}^{\mathrm{T}}_{sub,j,l,i}, \hat{\bar{x}}^{\mathrm{T}}_{sub,j,l,i}]^{\mathrm{T}}, \bar{u}_{f,j,i}, \hat{\bar{y}}_{\Theta_{f,dis,i},i}, \bar{\bar{y}}_{\Theta_{f,dis,i},i})$ and $\|\tilde{e}_{j,i,k}(t_k)\| \leq \hat{K}_{\tilde{e}_{j,i}}$.

Now we need to determine the supremum for $\|x_{sub,j,l,i}(t_{k+1}) - \hat{x}_{sub,j,l,i}(t_{k+1})\|$. For $t \in [t_k, t_{k+1})$, the scaled estimation error corresponding to jth residual of the ith LFDI scheme evolves as follows:

$$\dot{e^{j,i}} = \frac{1}{\varepsilon^{j,i}} A_0^{j,i} e^{j,i} + B^{j,i}(\phi^{j,i}(\bar{u}_{f_i}) - \phi_0^{j,i}(\bar{u}_{f_i})) + \eta_{j,i} \tag{4.19}$$

Using the same line of argument as the proof of Lemma 3.1 in Chapter 3, we get:

$$\|x_{sub,j,l,i} - \hat{x}_{sub,j,l,i}\| \leq L_{2,sub,,j,l,i}\beta_3(\varepsilon_{l,i})E_{s,sub,j,l,i} + \bar{\theta}_{sub,j,i}\Delta \tag{4.20}$$

where $E_{s,sub,i,l,j} = e^{-\alpha_{sub,i,l,j}\Delta}\|e^i_{sub,j,l,i}(t_k)\| + K_{B,sub,j,l,i}K_{\phi,,sub,j,l,i}\frac{1-e^{-\alpha_{sub,j,l,i}\Delta}}{\alpha_{sub,j,l,i}}$
$+ \frac{1-e^{-\alpha_{sub,j,l,i}\Delta}}{\alpha_{sub,j,l,i}}\bar{\eta}_{sub,j,l,i}$, $\alpha_{j,i}$, $K_{B_{j,i}}$, $K_{\phi_{j,i}}$ are the spectrum bound of the matrices $\frac{A_0^{j,i}}{\varepsilon_{l,i}}$, $B^{j,i}$, $(\phi^{j,i}(\zeta_{j,i}, \bar{u}_{f_{j,i}}, t) - \phi_0^{j,i}(\hat{\zeta}j, i, \bar{u}_{f_{j,i}}, t))$ respectively, $e_b^{*,j,i}$ is the upper bound on $\|e^{j,i}(t_k)\|$, $\beta_3(\varepsilon_{l,i}) = \max\{1, \varepsilon_{l,i}^{\omega_{\max}-1}\}$ and $L_{2,l,i} > 0$ is Lipschitz constant that satisfies the following inequality:

$$\|x_{j,i} - \hat{x}_{j,i}\| = \|T_{l,i}^{-1}(\zeta_{l,i}, u) - T_{l,i}'^{-1}(\hat{\zeta}_{l,i}, u_{l,i})\| \leq L_{2,l,i}\beta_3(\varepsilon_{l,i})E_{s,j,l,i} + \bar{\theta}_{j,i}\Delta \tag{4.21}$$

Using Eqs. 4.13, 4.18 and 4.20, for $\tilde{t}_{k'} \leq t_k \leq t_{f_k,dis}$, $r_{j,l,i,k+1}$ is bounded as below:

$$r_{j,l,i,k+1} \leq L_{2,sub,j,l,i}\beta_3^i(\varepsilon_{l,i})E_{s,sub,j,l,i} + \bar{\theta}_{sub,j,l,i}\Delta + E_p^{j,i} \tag{4.22}$$

where $E_p^{j,i} = (M^{j,i} + M'^{j,i})\Delta + K_{\tilde{e}}^{j,i}$ and for $t_k > t_{f_k,dis}$,

$$r_{j,l,i,k+1} \leq L_{2,sub,j,l,i}\beta_3^i(\varepsilon_{l,i})E_{s,sub,j,l,i} + \bar{\theta}_{sub,j,l,i}\Delta + \hat{E}_p^{j,i} \tag{4.23}$$

where $\hat{E}_p^{j,i} = (M^{j,i} + \hat{M}'^{j,i})\Delta + \hat{K}_{\tilde{e}}^{j,i}$. This concludes proof of Lemma 4.1. □

We select the threshold corresponding to each residual as below:

$$\delta_{j,i} = \begin{cases} L_{2,sub,j,l,i}\beta_3^{j,i}(\varepsilon_{j,i})E_{s,sub,j,l,i} + \bar{\theta}_{sub,j,i}\Delta + E_p^{j,i}, t_{k'} \leq t_k \leq t_{f_k,dis,i} \\ L_{2,sub,j,l,i}\beta_3^i(\varepsilon_{l,i})E_{s,sub,j,l,i} + \bar{\theta}_{sub,j,l,i}\Delta + \hat{E}_p^{j,i}, t_k > t_{f_k,dis,i} \end{cases} \tag{4.24}$$

Thus, it follows from Lemma 1 that under fault free condition $r_{j,i,k+1} \leq \delta_{j,i}$ i.e., there is no false alarm before occurrence of local faults in the ith subsystem.

Remark 4.6. Note that in the case of network systems with relatively low degree of nonlinearity like interconnected inverted pendulums mounted on carts of Zhang and Zhang (2012) or series of tanks of Ferrari *et al.* (2012) and Boem *et al.* (2017), the parameters in Eq. 4.24 can be specified and as a result, the Eq. 4.24 can be used directly for defining thresholds. However, when it comes to highly nonlinear networked systems like the reactor-separator example used here, it is not possible to find all of the constants in the Eq. 4.24. Instead we used simulations to determine the suprema of $\|\tilde{x}_{sub,j,l,i}(t_{k+1}) - x_{sub,j,l,i}(t_{k+1})\|$ and $\|x_{sub,j,l,i}(t_{k+1}) - \hat{x}_{sub,j,l,i}(t_{k+1})\|$, to in turn utilize as the threshold values suggested by Eq. 4.24. Note that the main

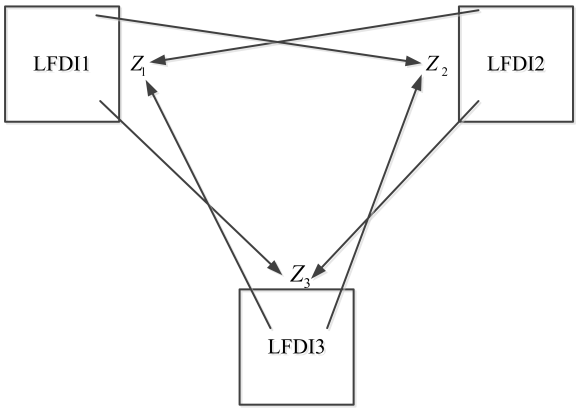


Figure 4.1: Schematic of the proposed distributed FDI framework for a network with three subsystems. LFDI schemes communicate to exchange information. This results in retaining detectability and isolability properties of LFDI schemes in the presence of fault in the shared interconnections between subsystems of the network.

purpose of deriving a mathematical formula for thresholds (Eq. 4.24) in this work is to rigorously establish the ability of the proposed distributed framework to achieve fault detection and isolation in the presence of uncertainty and fault in the shared interconnections of the network.

Note that the ith FDI filter receives the following information from the neighboring LFDI schemes:

$$Z_i = \begin{cases} \tilde{y}_i, & t_{k'} \leq t_k \leq t_{f_k,dis,i} \\ \hat{\tilde{y}}_i, & t_{f_k,dis,i} < t_k \end{cases} \tag{4.25}$$

where $\hat{\tilde{y}}_i = [\hat{\tilde{y}}_{\Theta_{f,dis,i},i}, \bar{\tilde{y}}_{\Theta_{f,dis,i},i}]^T$.

Figure 4.1 shows a schematic of the proposed distributed FDI framework. A fault in the ith subsystem is detected when at least one of the residuals in the ith LFDI scheme breaches its thresholds i.e. $r_{j,i} > \delta_{j,i}$ for some j in the LFDI scheme. We denote time of fault detection as $t_{d,i}$. Corollary 4.1 establishes that a residual designed to be insensitive to a specific fault scenario (using the existing approach in

Du *et al.* (2013)) retains its insensitive property even in the presence of uncertainty when thresholds are chosen using the proposed approach. To this end, let $r_{ins}^{j,i}$ denote the vector of residuals in the ith LFDI scheme designed to be insensitive to the jth fault scenario $\Theta_{f_{j,i}}$ in the ith scheme. Note that the proof is omitted here since it follows the same line of arguments as the results for FDI in the presence of uncertainty in Chapter 3.

Corollary 4.1. *Consider the ith subsystem of networked system described by Eq. 4.1, for which Assumptions 4.1- 4.6 hold and the fault detection and isolation framework characterized by residuals and thresholds described by Eq. 4.10 and Eq. 4.24, respectively and with* $\Theta_{f,j,i}(t) \neq 0$*, then* $r_o \in r_{ins}^{j,i}$*,* $r_o \leq \delta_{j,i}\ \forall\ t > t_{f,i} \geq t_{k'}$*.*

Detectability Analysis for Simultaneous Actuator and Sensor Faults

Fault detection only happens when at least one of the sensitive residuals to a fault scenario breaches its thresholds. However, in the presence of uncertainty, depending on fault functionality, the fault effect can be compensated by uncertainty and as a result, the corresponding sensitive residual does not breach its threshold (see e.g., Frank (1990); Chen and Patton (2012); Isermann (2006); Zhang *et al.* (2010b); Blanke *et al.* (2006); Ding (2008)). Thus it is required to establish the conditions necessary and sufficient for the residuals designed to be sensitive to a particular fault to breach their thresholds in the presence of uncertainty via a detectability analysis. Theorem 4.2 presents the sufficient conditions for simultaneous single actuator and single sensor faults to be detectable by the proposed FDI framework. The proof of Theorem 4.2 is similar to the proof for the centralized FDI methodology presented in Chapter 3 and hence is omitted here. Let $r_{\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}$ denote a sensitive residual to simultaneous faults

$u_{f_{j,i}}$ and $y_{f_{j,i}}$ defined as:

$$r_{\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}},k+1} = \|\tilde{x}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}(t_{k+1}) - \hat{x}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}(t_{k+1})\| \tag{4.26}$$

Theorem 4.2. *Consider the ith subsystem of the networked system described by Eq. 4.1, for which Assumptions 4.1-4.6 hold and the fault detection and isolation framework characterized by residual and threshold described by Eq. 4.10 and Eq. 4.24, respectively, and that a single actuator $u_{f_{j,i}}$ and single sensor fault $y_{f_{j,i}}$ occur simultaneously at time $t_{f,i}$: If there exists an interval of time $[t_f, t_d]$ where $t_f \geq t_{k'}$, such that the fault functions $u_{f_{j,i}}$ and $y_{f_{j,i}}$ satisfy*

$$\begin{aligned}
&\|\frac{1}{L'_{2,\bar{u}_{f_{j,i}}}\beta_3^{\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}(\varepsilon_{l,i})}\|\vartheta(\frac{A_{0,sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}}{\varepsilon_{l,i}}, H_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}, y_{f_{j,i}})\| \\
&+ \|Dev_{t_f\ldots t_d}(\tilde{x}^{\mathrm{T}}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}, \bar{u}_{f_{j,i}}, f_{d_{l,i}}, x^{\mathrm{T}}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}, u_{f_{j,i}})\| \\
&+ \|f_{d,u_{f_{j,i}}}\| - \delta'_{\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}} - \delta_{\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}\| > \delta_{\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}
\end{aligned} \tag{4.27}$$

where $Dev_{t_f\ldots t_d}(\tilde{x}^{\mathrm{T}}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}, \bar{u}_{f_{j,i}}, f_{d_{l,i}}, x^{\mathrm{T}}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}, u_{f_{j,i}})$
$= \int_{t_f}^{t_{f+1}}(dev(\tilde{x}^{\mathrm{T}}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_i}}, \bar{u}_{f,i}, f_{d_j}) - G_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}([x^{\mathrm{T}}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}, \bar{x}^{\mathrm{T}}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}]^{\mathrm{T}})u_{f_{j,i}})d\tau + \cdots +$
$\int_{t_{d-1}}^{t_d}(dev(\tilde{x}^{\mathrm{T}}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}, \bar{u}_{f,i}, f_{d_{l,i}}) - G_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}([x^{\mathrm{T}}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}, \bar{x}^{\mathrm{T}}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}]^{\mathrm{T}})u_{f_{j,i}})d\tau,$
$dev(\tilde{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_{j,i}}}, \tilde{x}^{\mathrm{T}}_{sub,u_{f_{j,i}},\bar{y}_{f_{j,i}}}, u_{sub,i}, f_{d_{l,i}}) = \tilde{f}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}([\tilde{x}^{\mathrm{T}}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}]^{\mathrm{T}}) +$
$\tilde{G}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}([\tilde{x}^{\mathrm{T}}_{sub,\bar{u}_{f_i},\bar{y}_{f_i}}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}]^{\mathrm{T}})u_{sub,j,i}$
$+ \tilde{I}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}},t}([[\tilde{x}^{\mathrm{T}}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}]^{\mathrm{T}}, \bar{u}_{f,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}, \tilde{y}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}})$
$- \tilde{f}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}([\tilde{x}^{\mathrm{T}}_{sub,u_{f_{j,i}},y_{f_{j,i}}}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,u_{f_{j,i}},y_{f_{j,i}}}]^{\mathrm{T}}) - \tilde{G}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}([\tilde{x}^{\mathrm{T}}_{sub,u_{f_{j,i}},y_{f_{j,i}}}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,u_{f_{j,i}},y_{f_{j,i}}}]^{\mathrm{T}})$
$u_{sub,j,i} - \tilde{I}_{sub,u_{f_{j,i}},y_{f_{j,i}},t}([[\tilde{x}^{\mathrm{T}}_{sub,u_{f_{j,i}},y_{f_{j,i}}}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,u_{f_{j,i}},y_{f_{j,i}}}]^{\mathrm{T}}, \bar{u}_{f,u_{f_{j,i}},y_{f_{j,i}}}, \tilde{y}_{sub,u_{f_{j,i}},y_{f_{j,i}}})$, *where* $u_{sub,j,i}$

is the subset of inputs corresponding to $G_{sub,\bar{u}_{f_{i,j}},\bar{y}_{f_{i,j}}}$, *where* l,i *refers to the corresponding observer used for defining* $r_{\bar{u}_{f_{i,j}},\bar{y}_{f_{i,j}}}$, *where*

$$\tilde{I}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}},t} = \begin{cases} \tilde{I}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}([\tilde{x}^{\mathrm{T}}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}]^{\mathrm{T}}, \bar{u}_{f,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}, \\ \tilde{y}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}), t_f \leq t_{f_k,dis} \\ \tilde{I}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}([\tilde{x}^{\mathrm{T}}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}, \hat{\tilde{x}}^{\mathrm{T}}_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}]^{\mathrm{T}}, \bar{u}_{f,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}, \hat{\tilde{y}}_{\Theta_{f,dis,i},\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}, \\ \bar{\tilde{y}}_{\Theta_{f,dis,i},\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}), t_f > t_{f_k,dis} \end{cases} \tag{4.28}$$

and $f_{d_{l,i}}$ *is the deviation of state estimates value from system states after fault occurrence:*

$$\begin{aligned} f_{d_{l,i}} &= \hat{x}_{\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}} - \hat{x}_{u_{f_{j,i}},y_{f_{j,i}}} = T'^{-1}_{l,i}(\hat{\zeta}_{\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}, u_{j,i}) - T'^{-1}_{l,i}(\hat{\zeta}_{faultfree,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}, u_{f_{j,i}} + u_{j,i}) \\ &= T'^{-1}_{l,i}(\hat{\zeta}_{\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}, u_{j,i}) - T'^{-1}_{l,i}(\hat{\zeta}_{\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}, u_{f_{j,i}} + u_{j,i}) + T'^{-1}_{l,i}(\hat{\zeta}_{\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}, u_{f_{j,i}} + u_{j,i}) \\ &- T'^{-1}_{l,i}(\hat{\zeta}_{faultfree,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}, u_{f_{j,i}} + u_{j,i}) \end{aligned} \tag{4.29}$$

where $\hat{\zeta}_{\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}} = \hat{\zeta}_{faultfree,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}} + H_{\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}} \int_{t_f}^{t_{f+1}} e^{(A_{j,i} - H_{\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}} C)(t_{f+1}-\tau)} y_{f_{j,i}} d\tau + \ldots$
$+ H_{\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}} \int_{t_{d-1}}^{t_d} e^{(A_{j,i} - H_{\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}} C_{j,i})(t_d-\tau)} y_{f_{j,i}} d\tau$, $\vartheta(\frac{A_{0,sub,j,i}}{\varepsilon_{l,i}}, H_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}, y_{f_{j,i}}) =$
$\int_{t_f}^{t_{f+1}} e^{\frac{A_{0,sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}}{\varepsilon_{l,i}}(t_{f+1}-\tau)} [D_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}]^{-1} H_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}} y_{f_{j,i}} d\tau + \ldots$
$+ \int_{t_{d-1}}^{t_d} e^{\frac{A_{0,sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}}{\varepsilon_{l,i}}(t_d-\tau)} [D_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}]^{-1} H_{sub,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}} y_{f_{j,i}} d\tau,$
$f_{d,u_{f_{j,i}}} = T_j^{-1}(\zeta_{\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}, u_{f_i} + u_i) - T_j^{-1}(\zeta_{\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}, u_i)$ *and* $\delta'_{\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}} = E'_{s,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}} =$
$(\frac{1}{L'_{2,u_{j,i}} \beta_3^{\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}(\varepsilon_{l,i})} - L_{2,\bar{u}_{f_{j,i}}} \beta_3^{\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}(\varepsilon_{l,i})) E_{s,\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}$, *then the fault is detected i.e.,* $r_{\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}},d} > \delta_{\bar{u}_{f_{j,i}},\bar{y}_{f_{j,i}}}$.

Remark 4.7. Note that to check the detectability condition presented in Theorem 2, we simply calculate the infimum of $\|\|x_{sub,j,l,i}(t_{d_i}) - \hat{x}_{sub,j,l,i}(t_{d_i})\| - \|\tilde{x}_{sub,j,l,i}(t_{d_i}) - x_{sub,j,l,i}(t_{d_i})\|\|$ for each residual that is called the detectability constant, $\bar{\delta}_{j,i}$ (see the

results presented in Chapter 3). If the detectability constant is more than the value of its corresponding threshold then the residual is expected to breach the threshold (see Chapter 3 for more on this).

Isolability Condition

Having presented the detectability condition corresponding to different faulty scenarios in the LFDI scheme corresponding to the ith subsystem, Theorem 4.3 presents the fault isolation logic for the identification of faulty component in the ith subsystem that also serves as the FDI mechanism. The proof of Theorem 4.3 follows a similar line of as results presented in the literature (see e.g., Du *et al.* (2013); Shahnazari *et al.* (2016)) and hence is omitted here.

Theorem 4.3. *Consider the ith subsystem of the network presented by Eq. 4.1, for which Assumptions 4.1-4.6 hold. If $r_{j,i,t_d} > \delta_{j,i}$, for all $j \in \{1, \ldots, n_{f,dis_i}\} \backslash w$ then*

$$\begin{cases} \Theta_{f,w,i}(t) \neq 0 \quad or \quad \Theta_{f,dis,i}(t) \neq 0, \quad t_d \leq t < t_{f,k,dis,i} \\ \Theta_{f,w,i}(t) \neq 0 \quad and \quad \Theta_{f,dis,i}(t) \neq 0, \quad t_{f,k,dis,i} \leq t_d \leq t \end{cases} \tag{4.30}$$

for some $t \in [t_d, t_{d+1})$.

Remark 4.8. Note that Theorem 4.3 presents the conditions under which a fault can be isolated in networked systems. These conditions are affected by the network structure and the existing uncertainties in the system model. To understand this better, consider the case where the breaching pattern in the ith LFDI schemes matches the signature of a multiple fault scenario. This means that only one residual does not breach its threshold. If the corresponding local subsystem is not affected by any shared interconnection from other subsystems, the fault is successfully isolated.

However, in the cases where the corresponding subsystem is affected by shared interconnections from other subsystems, there is always the possibility that the shared interconnections are faulty. If the shared interconnections are local variables in some other subsystem, and the local FDI scheme for that subsystem (or a smart sensor) is able to unequivocally provide fault information for the shared variable, then this issue can be resolved. On the other hand, if clear information about the shared interconnection is not available, then no definitive conclusions can be drawn. Theorem 3 imposes a necessary condition on network structure under which each chain of subsystems affecting the ith subsystem must be a cascade to enable the corresponding LFDI scheme to isolate faults in the ith subsystem. Note that this is a fundamental limitation of the network caused by the inherent trade-off between robustness and fault sensitivity, not the proposed FDI methodology and this assumption is also utilized by other FDI frameworks (see e.g., Ferrari *et al.* (2012); Zhang and Zhang (2012); Reppa *et al.* (2015)).

Note that the thresholds are defined using the suprema (lowest upper bound). Thus, the probability of a sensitive residual breaching its threshold when expected is maximized. As a result of this, when a breaching pattern matches the signature of a fault scenario, if thresholds values are relatively small with respect to the acceptable range for residual values when assuming a uniform distribution for each fault scenario, the most likely source of fault is the corresponding fault scenario to the observed beaching pattern. Thus to handle situations where the conditions necessary for FDI are not satisfied, we present the fault detectability index that measures the probability of detecting a fault scenario by its corresponding sensitive residuals, defined as below:

$$P_{j,i} = \frac{sup(r_{j,i}) - \delta_{j,i}}{sup(r_{j,i})} \tag{4.31}$$

where $P_{j,i}$ can be understood as the probability of the jth residual of the ith LFDI scheme breaching its threshold when a fault scenario that belongs to $\bar{\Theta}_{j,i}$ occurs, $sup(r_{j,i})$ denotes the lowest upper bound possible for the jth residual corresponding of the ith LFDI scheme when a fault scenario that belongs to $\bar{\Theta}_{j,i}$ occurs. Note that the supremum of $r_{j,i}$ is calculated by finding the maximum possible values for the $r_{j,i}$ in the presence of each one of the fault scenarios that belongs to $\bar{\Theta}_{j,i}$ using an acceptable range for each fault scenario and selecting the lowest value of these maximums as supermimum of $r_{j,i}$. If a residual has detectability index higher than 50 %, we consider the corresponding residual not breaching its threshold trustworthy for being utilized in decision making.

Remark 4.9. Note that in this work it is assumed that the occurrence of different fault scenarios and the possible fault functionalities corresponding to each of these fault scenarios have uniform distributions. Note that this is a fair assumption, since the proposed framework is designed based on the idea that no prior knowledge is available regarding plant fault history. However, if such information is available, the distributions corresponding to the occurrence of fault scenarios and the fault functionalities can be obtained from the plant fault history data. In this case, the detectability index can be calculated using the probability density function (PDF) obtained from the plant data.

Having presented the FDI design for ith LFDI scheme rigorously, Algorithm 4.1 summarizes the calculations and the decision making procedure of the ith LFDI scheme at kth sampling time.

Algorithm 4.1. 1. Initialize the state estimators and state predictors using state estimates at $k-1$th sampling time, $\hat{x}_{l,i}(k)$ where $l = 1, \ldots, p_i$.

2. If $\Theta_{f,dis,i} \neq 0$ i.e., a fault in one of the shared interconnections with the ith subsystem is diagnosed, then the ith LFDI scheme replaces faulty measurements by the healthy estimations provided by the corresponding LFDI scheme to $\tilde{y}_{i,\Theta_{f,dis,i}}$ and updates the thresholds corresponding to the ith scheme using Eq. 4.24 accordingly.

3. Compute values of state estimates $\hat{x}_{l,i}(k)$ where $l = 1, \ldots, p_i$, state prediction $\tilde{x}_{j,i}(k)$ where $j = 1, \ldots, n_{f,i}$ and residuals $r_{j,i}$ where $j = 1, \ldots, n_{f,i}$.

4. If $r_{j,i} > \delta_{j,i}$, a fault in the ith FDI scheme is detected.

5. If $r_{j,i} > \delta_{j,i}$ for all $j \in \{1, \ldots, n_{f_i}\} \backslash w$ and $P_{w,i} > 0.5$, then the ith LFDI scheme claims the corresponding fault scenario to $\Theta_{f,w,i}(t)$ has occurred.
   Otherwise, for $t < t_{f,k,dis,i}$, the ith LFDI scheme notifies that at least one of the fault scenarios corresponding to $\Theta_{f,w,i}(t)$ and $\Theta_{f,dis,i}(t)$ has occurred, and for $t_{f,k,dis,i} \leq t_d \leq t$, the ith LFDI scheme declares both of the fault scenarios corresponding to $\Theta_{f,w,i}(t)$ and $\Theta_{f,dis,i}(t)$ have occurred.

Remark 4.10. Note that the main advantage of this work with respect to the existing results in the literature for networked systems (see e.g., Zhang and Zhang (2012); Ferrari *et al.* (2012); Keliris *et al.* (2015); Peng *et al.* (2015); Reppa *et al.* (2015); Yin and Liu (2017)) is to diagnose simultaneous faults in nonlinear uncertain systems when the shared interconnections are faulty. Also, the proposed distributed FDI methodology, unlike some of the existing results in the literature (see e.g.,Ferrari

et al. (2012); Reppa *et al.* (2015)) does not require a global fault diagnoser to make decision in the presence of fault in the shared interconnections of the network. Each LFDI scheme can make a decision using the available measurements and information provided by the neighboring LFDI schemes. This results in increased reliability of the proposed distributed FDI scheme.

4.5 Simulation example

This section illustrates application the proposed FDI methodology to a network of three vessels, reactor - separator process as shown in Figure 3 (see e.g., Peng *et al.* (2015)). The process variables are defined in Table 4.1 (for information regarding reactions taking place in the plant, the process model and process parameters see e.g., Peng *et al.* (2015)).

The control objective is to stabilize the plant at unstable steady state point $y_1 = C_{\mathrm{A1}} = 3.31$ kmol/m^3, $y_2 = C_{\mathrm{B1}} = 0.17$ kmol/m^3, $y_3 = C_{\mathrm{C1}} = 0.04$ kmol/m^3, $y_4 = T_1 = 369.5$ K3, $y_5 = C_{\mathrm{A2}} = 2.75$ kmol/m^3, $y_6 = C_{\mathrm{B2}} = 0.45$ kmol/m^3, $y_7 = C_{\mathrm{C2}} = 0.11$ kmol/m^3, $y_8 = T_2 = 435.2$ K, $y_9 = C_{\mathrm{A3}} = 2.88$ kmol/m^3, $y_{10} = C_{\mathrm{B3}} = 0.5$ kmol/m^3, $y_{11} = C_{\mathrm{C3}} = 0.12$ kmol/m^3, $y_{12} = T_3 = 435.2$ K. The manipulated input variables are $u = [Q_1, Q_2, Q_3]^{\mathrm{T}}$, where $\|u_1\| \leq 5 \times 10^4$ kJ/hr, $\|u_2\| \leq 1.5 \times 10^5$ kJ/hr and $\|u_3\| \leq 2 \times 10^5$ kJ/hr. It is assumed that all of the states are measurable. A local robust Lyapunov based model predictive controller is designed for each subsystem using the Lyapunov-based MPC design of Mahmood *et al.* (2008a).

The hold time for control action is selected $\Delta = 0.01$ hr for all three controllers. The weighting matrices used to penalize the deviations of the state and input from their nominal values for the ith local controller are chosen as $Q_{w_i} =$

Table 4.1: Definition of the process variables used for the network of chemical reactor example used in this work.

Parameter	Definition
C_{Aj0}	Concentration of A in the feed stream to tank j, $j = 1, 2$
C_{ij}	Concentration of species i, $i = A, B, C$ in tank j, $j = 1, 2, 3$
T_{j0}	Temperature of the feed stream to tank j, $j = 1, 2$
F_{j0}	Flow rate of the feed stream to tank j, $j = 1, 2$
F_j	Flow rate of the effluent stream from tank
H_{vap}	Heat of vaporization
Q_i	Heat input to tank j, $j = 1, 2, 3$

$diag[10^3, 10^3, 10^3, 20]$ and $R_{w_i} = 10^{-12}$, respectively, where $i = 1, ..3$. The Lyapunov function for the ith subsystem is chosen as $V(x_i) = \bar{x}_i' P_i \bar{x}_i$ where $\bar{x}_i = x_{i,n} - x_i$ is the vector of deviation variables, $x_{i,n}$ denotes the vector desired nominal values of the states of the ith subsystem and $P_1 = diag[10^3, 10^3, 10^3, 2 \times 10^2]$, $P_2 = diag[10^3, 10^3, 10^3, 2 \times 10^4]$ and $P_3 = diag[10^3, 10^3, 10^3, 2 \times 10^2]$.

Each subsystem of the network is subject to modeling uncertainty and measurement noise. In particular, the values of C_{A20} and H_{vap} are 5% less than their nominal values. Furthermore, inlet temperature to the tank 1, T_{10} fluctuates with time, with the actual flow rate being $1 + 0.05\sin(t)$ times of its nominal value. The known bounds on this uncertainties are 10%, 10%, and 5 % of their nominal values. The concentration and temperature measurements have combinations of 5 Hz sinusoidal noises. The magnitudes of the measurement noise over each sampling time follow a normal distribution with the standard deviations being 0.01 kmol/m^3 and 0.1 K for concentrations and temperatures, respectively. The noisy measurements are passed through a first-order low-pass filter with the filter time constant being 3.6 seconds.

Note that the proposed framework in this work only accounts for fault scenarios

that are distinguishable locally. Thus it is assumed local indistinguishable fault scenarios do not take place. For this sake, it is assumed none of $y_2 = C_{B1}$, $y_3 = C_{C1}$, $y_6 = C_{B2}$, $y_7 = C_{C2}$ and $y_{12} = T_3$ are faulty, since these outputs are not observable and as a result of this, fault in none of the corresponding sensors to this outputs can be isolated (see the results in Chapter 2 and Shahnazari *et al.* (2016) for more on this).

A bank of observers is required for designing LFDI scheme corresponding to each subsystem. To this end, based on the above explanation, three observers are designed to estimate states of the each subsystem, that results in a total number of nine observers. For the LFDI schemes corresponding to the first and second subsystems, two observers are designed by using three of the outputs while the third one is designed using only two of the outputs. For the LFDI scheme corresponding to the third subsystem, all of the three observers are designed using only three of the available measurements. Based on the methodology presented in this work, 6, 6 and 7 residuals are generated for the first, second and third LFDI schemes, respectively. The thresholds are selected based on Eq. 4.13 via simulations. To this end, the summation of the maximum observed values for $\|\tilde{x}_{sub,j,l,i}(t_{k+1}) - x_{sub,j,l,i}(t_{k+1})\|$ and $\|x_{sub,j,l,i}(t_{k+1}) - \hat{x}_{sub,j,l,i}(t_{k+1})\|$ by considering all possible combinations of the bounds on uncertainties. A value slightly larger is selected as the corresponding threshold for each residual, and reported in Tables 4.2, 4.3 and 4.4. Also, the thresholds corresponding to the second and third LFDI schemes in the case of distributed faults in $y_1 = C_{A1}$,$y_4 = T_1$, $y_5 = C_{A2}$ and $y_8 = T_2$ are reported in second column of threshold tab in Tables 4.3 and 4.4.

We next consider a case where simultaneous faults take place in $y_1 = C_{A1}$ and

Table 4.2: Faults to which the residuals of the FDI scheme corresponding to the first subsystem are insensitive and thresholds for the fault isolation design for the case study based on the proposed framework.

Residual	Faults	Threshold	Residual	Faults	Threshold
$r_{1,1}$	y_{f_1}	0.2	$r_{1,2}$	y_{f_4}	1.2
$r_{1,3}$	y_{f_1}, y_{f_4}	0.07	$r_{1,4}$	u_{f_1}	0.04
$r_{1,5}$	y_{f_4}, u_{f_1}	0.04	$r_{1,6}$	y_{f_1}, u_{f_1}	0.2

Table 4.3: Faults to which the residuals of the FDI scheme corresponding to the second subsystem are insensitive and thresholds for the fault isolation design of the case study based on the proposed framework.

Residual	Faults	Threshold		Residual	Faults	Threshold	
$r_{2,1}$	y_{f_5}	1.03	1.16	$r_{2,2}$	y_{f_8}	1.33	1.46
$r_{2,3}$	y_{f_5}, y_{f_8}	2.62	2.71	$r_{2,4}$	u_{f_2}	0.13	0.13
$r_{2,5}$	y_{f_8}, u_{f_2}	0.13	0.14	$r_{2,6}$	y_{f_5}, u_{f_2}	0.13	0.14

$y_4 = T_1$ with functionalities of $(-1.44 - 0.1 \sin t)(1 - e^{10(t_{f1}-t)})$ and -15, respectively, at time $t_{f1} = 0.5$ hr. This is followed by simultaneous faults in $u_2 = Q_2$ and $y_8 = T_2$ with functionalities of 27500 and 50, respectively, at time $t_{f2} = 1$ hr. In the first LFDI scheme, all of the residuals breach their thresholds except $r_{1,3}$ that matches the signature of simultaneous faults in $y_1 = C_{A1}$ and $y_4 = T_1$. But, since the first subsystem is affected by the shared interconnections from the other subsystems, there is always the possibility that the shared interconnection is faulty and the insensitive residual is not breaching its threshold due to trade-off between robustness and uncertainty. Thus fault isolation can not be achieved. To address this, we compute the detectability index.

Note that the calculated detectability index values for all of the residuals are more than 70%, by considering ± 50 % of the acceptable range for each variable as the maximum possible size for faults taking place in the corresponding actuators and

Table 4.4: Faults to which the residuals of the FDI scheme corresponding to the third subsystem are insensitive and thresholds for the fault isolation design of the case study based on the proposed framework.

Residual	Faults	Threshold		Residual	Faults	Threshold	
$r_{3,1}$	y_{f_9}	0.25	0.25	$r_{3,2}$	$y_{f_{10}}$	0.04	0.06
$r_{3,3}$	$y_{f_{11}}$	0.02	0.05	$r_{3,4}$	u_{f_3}	0.25	0.25
$r_{3,5}$	y_{f_9}, u_{f_3}	0.25	0.25	$r_{3,6}$	$y_{f_{10}}, u_{f_3}$	0.05	0.05
$r_{3,7}$	$y_{f_{11}}, u_{f_3}$	0.02	0.01				

sensors. Thus we assume when a residual does not breach its threshold, it means the residual is most likely insensitive to that fault scenario. Note that the detectability index value for each threshold depends on the upper bounds utilized for uncertainties and the range considered for each fault scenario. Thus, using different values for the upper bounds of uncertainties or different range for each fault scenario result in different values for detectability index corresponding to each residual. By using the detectability index proposed in this work, since all of the residuals have detectability index higher than 70%, when a residual does not breach its threshold, it can be considered that not breaching is not due to trade-off with respect to uncertainty. As a result of this, the residual being insensitive is trustworthy for being utilized in decision making. Thus simultaneous faults in $y_1 = C_{A1}$ and $y_4 = T_1$ are successfully isolated.

In the second LFDI scheme, by using a decentralized FDI architecture, all of the residuals breach their thresholds at most in 10 minutes after fault occurrence in the first subsystem (the residuals profile is not presented here for the sake of brevity). Thus, the second LFDI issues a fault alarm at $t = 0.6$ hr. However, since of all the residuals have breached their thresholds, the second LFDI scheme can only act as detection scheme and fault isolation is not possible. Figure 4.2 shows the evolution of

residuals profile corresponding to the second LFDI scheme when using the proposed distributed FDI framework in this work. In this case, upon isolation of faults in the first LFDI scheme at $t = 0.6$ hr, the first LFDI scheme notifies the second LFDI scheme and transmits the healthy estimation of the $y_1 = C_{A1}$ and $y_4 = T_1$ to the second LFDI scheme. The second LFDI scheme replaces the faulty measurements of $y_1 = C_{A1}$ and $y_4 = T_1$ with the healthy estimation provided and updates the thresholds accordingly. As a result of this, all of the residuals breach their thresholds except $r_{2,5}$ that matches the signature of fault in $u_2 = Q_2$ and $y_8 = T_2$. Again without using the concept of detectability index proposed in this work, fault isolation can not be achieved and there is always the possibility that the shared interconnections affecting the second subsystems are faulty. However, as described before, since all of the residuals have detectability index higher than 70 %, when a residual does not breach its threshold, it can be considered that not breaching is not due to trade-off with respect to uncertainty. Thus simultaneous faults in $u_2 = Q_2$ and $y_8 = T_2$ are successfully isolated.

In the third LFDI scheme, using a decentralized architecture, some of the residuals breach their threshold (the results are not presented here for the sake of brevity). This results in fault detection. However, none of the actuators and sensors of the third subsystem are subject to fault. Using the distributed architecture proposed in this work, the second LFDI scheme notifies the third LFDI scheme and transmits the healthy estimation of the $y_8 = T_2$ to the third LFDI scheme, upon isolation of faults in the second LFDI scheme at $t = 1.09$ hr. The third LFDI scheme replaces the faulty measurements of $y_8 = T_2$ with the healthy estimation provided and updates the thresholds accordingly. This results in fast recovery of the third LFDI scheme

filters and removing the fault alarm by third LFDI scheme in 10 minutes (the results are not presented here for sake of brevity). Again without utilizing the concept of detectability index, there is always possibility that a fault has occurred in the third subsystem, but it cannot be isolated due to the trade-off between robustness and uncertainty. Figure 4.3 shows evolution of $r_{3,3}$ using a decentralized FDI framework and using the proposed distributed FDI methodology in this work. As can be seen, using the decentralized FDI scheme, $r_{3,3}$ breaches its thresholds, leading to a false decision making by the corresponding LFDI. However, utilizing the distributed FDI scheme, $r_{3,3}$ only breaches its threshold for a short period of time (less than 10 minutes), then recovers quickly as a result of using healthy estimation of the faulty shared interconnection and the updated value for thresholds. Note that in this case, using the updated value for threshold corresponding to $r_{3,3}$ results in quicker recovery of the FDI filter.

4.6 Conclusions

In this work, we addressed the problem of simultaneous fault diagnosis in nonlinear uncertain networked systems utilizing a distributed fault detection and fault isolation strategy. The idea is to design a bank of local robust FDI schemes in a distributed manner with each FDI scheme corresponding to a subsystem. Time-varying thresholds were selected by explicitly accounting for the effect of uncertainties and faults in shared interconnections. In this way, robustness of the LFDI schemes to false alarms is guaranteed. Also, in the case of faults in the shared interconnections that can be isolated locally, the distributed architecture of the proposed FDI framework allows the other FDI schemes to function as intended. The detectability and isoability

conditions were rigorously derived for the distributed FDI scheme. Effectiveness of the proposed methodology was shown via application to a reactor-separator process subject to uncertainty and measurement noise.

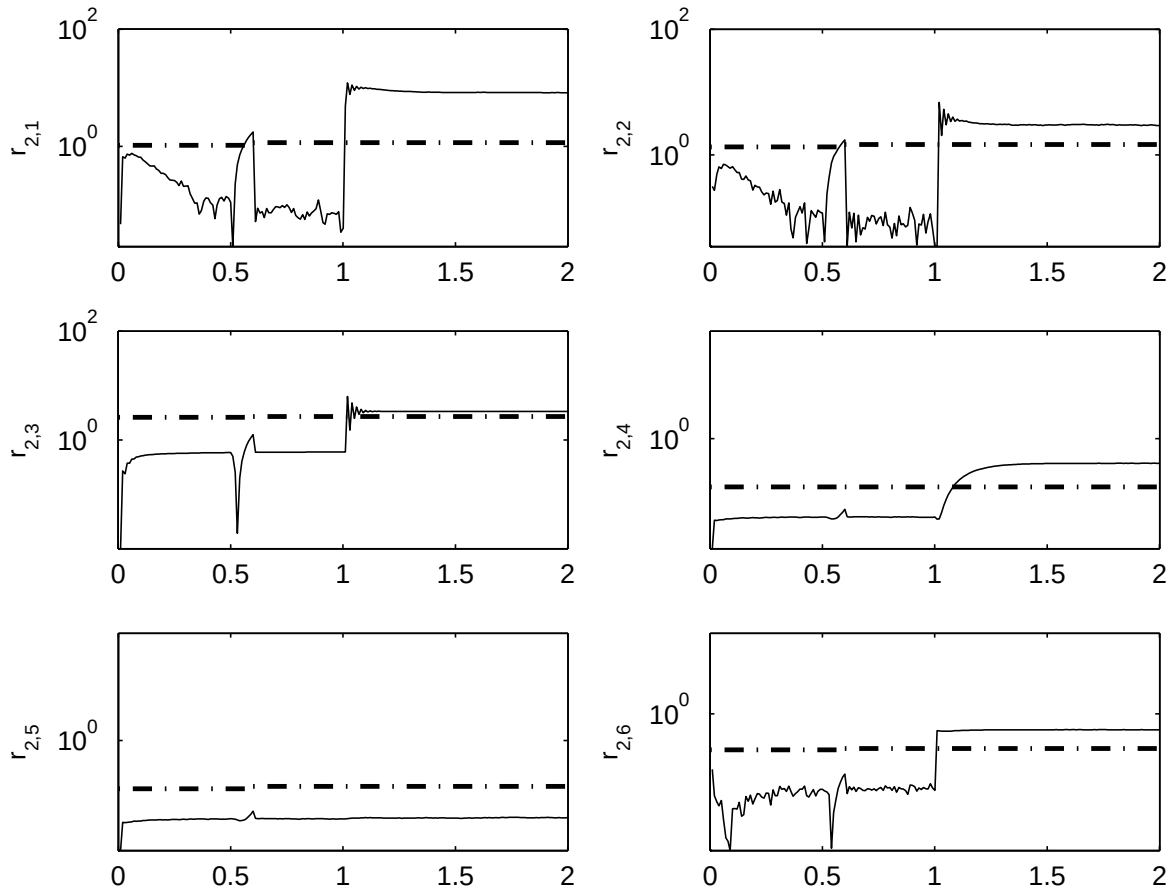


Figure 4.2: a) Evolution of the residual corresponding to the second LFDI scheme. As can be seen, by using the healthy estimates of $y_1 = C_{A1}$ and $y_4 = T_1$ upon isolation of fault in the first LFDI scheme at $t = 0.6$ hr, all of the residuals remain insensitive until occurrence of fault in $u_1 = Q_1$ and $y_8 = T_2$. Then all of the residuals breach their thresholds except $r_{2,5}$ that matches the signature of simultaneous faults in $u_1 = Q_1$ and $y_8 = T_2$ and as a result the fault is successfully isolated.

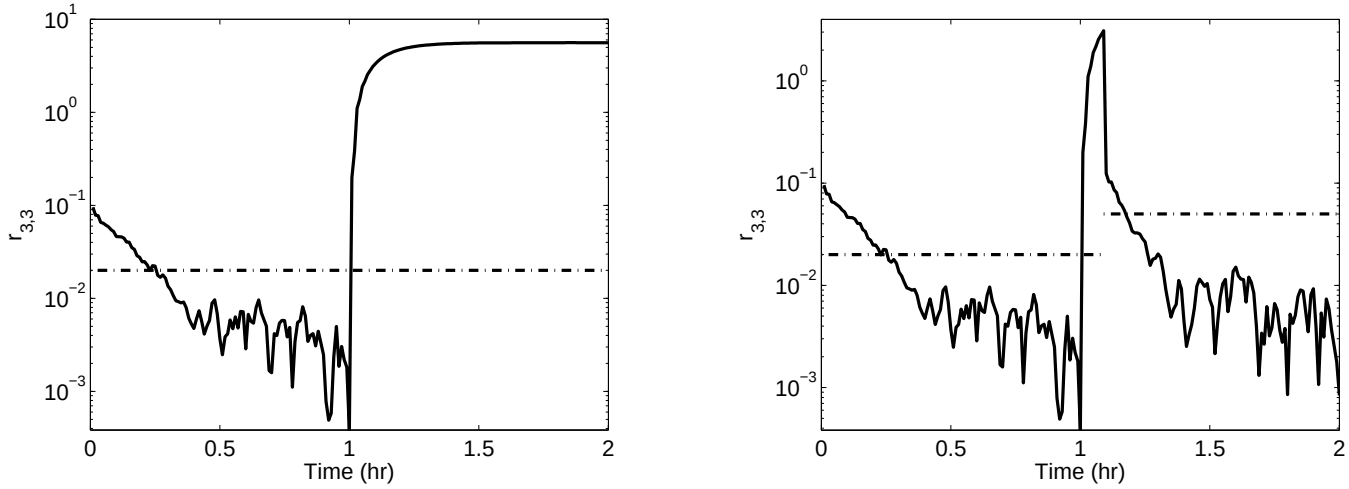


Figure 4.3: a) Evolution of the residual $r_{3,3}$ (solid lines) and thresholds (dashed-dotted lines) using a decentralized FDI framework. In this case, the residual breaches its threshold that results in a false decision by the LFDI scheme. b) Evolution of the residual $r_{3,3}$ (solid lines) and thresholds (dashed-dotted lines) using the distributed framework proposed in this work. In this case, the healthy estimations of $y_8 = T_2$ provided by the second LFDI scheme upon isolation of the fault at time $t = 1.09$ are utilized by the third LFDI scheme and thresholds are updated accordingly. As can be seen the residual recovers quickly and it does not breach its threshold as expected. This results in correct decision making by the corresponding LFDI scheme. Note that in this case, using the updated value for threshold corresponding to $r_{3,3}$ results in quicker recovery of the FDI filter.

Chapter 5

Heating, Ventilation and Air Conditioning Systems: Fault Detection and Isolation and Safe Parking

The contributions of this chapter have been submitted to/ published in:

Journal Papers:

Shahnazari, H., Mhaskar, P., House, J. M., and Salsbury, T. I. (2018). Heating, ventilation and air conditioning systems: Fault detection and isolation and safe parking. *Computers & Chemical Engineering*, **108**, 139 – 151.

Refereed Conference Proceedings:

Hadi Shahnazari, Prashant Mhaskar, John M House, and Timothy I Salsbury. Fault diagnosis design for heating, ventilation and air conditioning systems. In *American*

Control Conference (ACC), 2018, submitted.

5.1 Introduction

In the previous chapters, some theoretical results were presented addressing fault diagnosis in the presence of nonlinearities, uncertainties, high dimensionality and in the absence of enough analytical redundancy in the system structure. The efficiency of the proposed methodologies was shown via simulation case studies. In this chapter, an integrated fault diagnosis and safe parking framework is presented for HVAC systems as an industrial complex system composed of at least 10 highly interactive components.

Government regulations and initiatives have placed a large emphasis on the reduction of energy consumption and increase in energy efficiency. Heating, ventilation, and air-conditioning (HVAC) systems are responsible for 40-50% of total building energy consumption, motivating research on energy efficient building control (see, e.g., Ma *et al.* (2012), Mendoza-Serrano and Chmielewski (2012), Mendoza-Serrano and Chmielewski (2014), Cole *et al.* (2013), Cole *et al.* (2014), Touretzky and Baldea (2014a) and Touretzky and Baldea (2014b)). It is estimated that in the U.S. alone (see e.g., Schein *et al.* (2006)), fault detection and isolation (FDI), and fault tolerant control methods could be capable of saving 10-40% of HVAC energy consumption.

These realizations have motivated significant research effort on devising FDI frameworks for HVAC systems with many studies focusing on air handling unit (AHUs) and variable air volume (VAV) boxes. Existing frameworks utilize a statistic based approach for the purpose of FDI. In House *et al.* (2001), a fault detection tool is proposed that uses a set of expert rules derived from mass and energy balances to detect faults in air handling units (AHUs). A subset of the expert rules which correspond to the current mode of operation are then evaluated to determine whether

a fault exists. In Chen and Lan (2010), a PCA based approach is used to extract the correlation of measured variables in a heating/cooling building system and reduce the dimension of the measured data. Square prediction error (SPE) statistic is then used to detect sensor faults in the system. Then, a sensor validity index (SVI) is employed to identify the faulty sensor and a reconstruction algorithm is presented to recover the correct data for the faulty sensor in accordance with the correlations among system variables. In Du and Jin (2007), a combination of principal component analysis (PCA) and joint angle analysis are used to detect and isolate multiple faults in AHUs with variable air volume (VAV) boxes.

In Schein and House (2003), a fault detection method is developed for application to variable-air-volume (VAV) boxes using control charts. In Yoshida *et al.* (2001), a recursive autoregressive exogenous algorithm is used to develop a dynamic FDD model that addresses single fault scenarios in VAV boxes. In Wang and Qin (2005), a strategy using PCA is developed for detecting and validating flow sensor faults. The fault is detected using both the T^2 statistic and SPE and isolated using the SPE contribution plot. In Qin and Wang (2005), a hybrid approach utilizing expert rules, performance indexes and statistical process control models is used to address single fault scenarios in VAV boxes. In Du *et al.* (2007), a combination of PCA and joint angle analysis is used to diagnose sensor faults in VAV boxes. In Wu and Sun (2011), a cross-level fault detection methodology is proposed based on energy flow in HVAC systems that detects faulty HVAC units instead of component faults by comparing the current flow energy consumption in the system with respect to its normal expected patterns. The existing results in the literature, however, consider only isolation of single fault scenarios in the VAV boxes and do not consider multiple

sensor faults or multiple actuator faults in the VAV boxes, in part due to the limitation of the underlying statistical based approaches (as demonstrated via simulations in the present work).

In the area of dynamic model based FDI, there is a large body of methods in the literature utilizing linear model based FDI design, and these approaches can be categorized into parity relation and diagnostic observer (see e.g., Frank (1990), Venkatasubramanian *et al.* (2003) and Magni and Mouyon (1994)). Note that these methodologies are equivalent when it comes to residuals generation and both use output estimation error for defining residuals (see e.g., Gertler and Monajemy (1995) and Yoon and MacGregor (2000)). However, these methods have not been utilized to detect and isolate actuator faults where the effect of the fault is compensated by the controller. Thus, the area of FDI using linear models in general, and applications to HVAC systems in particular, stands to gain from novel linear model based FDI design that achieve FDI for sensor and actuator faults (including those masked by the controller).

There also exist results on fault tolerant control (FTC) of HVAC systems. In Seem (2001), the control design compensates for the effect of faults as much as possible by switching between different control modes available in the air handling unit design. In Hao *et al.* (2005), single sensor faults are diagnosed and handled via sensor redundancy. In Talukdar and Patra (2010), a model based fault tolerant control strategy is developed for handling multiple stuck dampers in the VAV boxes of HVAC systems. Fault tolerant control is achieved by modifying the airflow through the healthy zones. This is based on the assumption that the overall HVAC system maintains a constant total air flow rate. Under this assumption, changing the amount of air flow

entering the healthy zones affects the amount of air flow rate entering faulty zones. This assumption, however breaks down in applications where the static pressure is held constant. In Bengea *et al.* (2015), the fault tolerant control design of the HVAC system is based on real time estimation of the fault magnitude, and determining MPC constraints (input constraints) based on those values.

These fault-tolerant control approaches, however, are all predicated on the idea of maintaining nominal operation as the only control objective before and after fault occurrence, which might simply be impossible, or expensive in case of certain faults. Recently, safe-parking based approaches for fault-tolerant control have been proposed (see e.g., Gandhi and Mhaskar (2008) and Du and Mhaskar (2011)) that upon fault detection, prescribe temporarily operating (or 'parking') the process at an appropriate operating point, instead of trying to maintain nominal operation. Various algorithms for safe-parking have been proposed focusing on stability/optimality of the overall operation. These ideas, however, have not been applied to HVAC systems. In summary the area of VAV control stands to benefit from implementations that can handle multiple actuator and sensor fault detection and isolation, and implement safe parking based approaches.

Motivated by the above considerations, in this work, we design and implement an integrated framework for fault diagnosis and safe parking of VAV boxes of HVAC systems. To compare with existing approaches, first, a statistical model based FDI scheme is designed using existing PCA and joint angle analysis based techniques. Then we design linear causal model based frameworks for detection and isolation of multiple actuators and multiple sensor faults. The linear model is identified using a subspace identification method applied to data from a detailed Modelica model of

an AHU with five VAV boxes. The linear model based approach is seen to possess superior fault-isolation capabilities. Finally, the problem of fault handling in the context of VAV boxes is addressed, recognizing that while in the present context, the faults are not safety critical in nature, they do present an opportunity for trading off between comfort and energy usage. Thus, a safe parking strategy is designed to handle stuck dampers and the resulting energy reduction demonstrated.

5.2 Preliminaries

In this section, we briefly review the air handling unit (AHU) model first, then describe the VAV box model.

5.2.1 Air handling unit model

An air handling unit usually comprises fans, heating and cooling coils, and dampers to achieve the supply air temperature set point. To serve as a simulation test bed, we use a detailed Modelica model of an AHU with five VAV boxes. The testbed AHU model has three dampers (outdoor, recirculation, and exhaust), cooling and heating coils with valves and temperature, pressure and flow sensors for monitoring and control. Each of the actuators in the AHU model is controlled using a single loop proportional integrator (PI) controller. The control objective is to provide supply air with a constant temperature (typically 55^0F) at the downstream of the supply fan. The supply fan is used to maintain the static pressure in the supply duct at a constant value. Figure 5.1 shows a schematic diagram of a typical AHU.

The testbed AHU system has four modes of operation used for controlling the

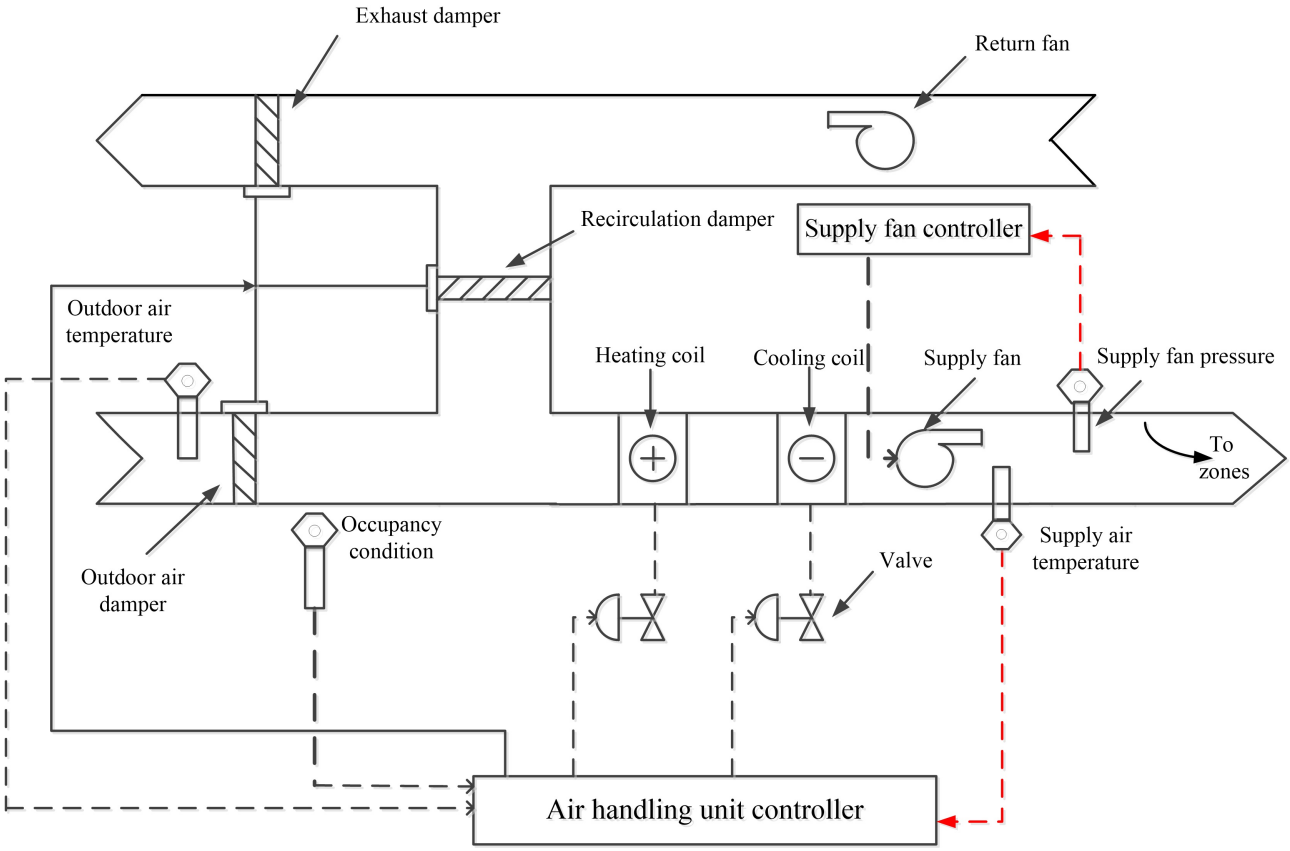


Figure 5.1: Schematic of an AHU

supply air temperature. A sequencing logic determines the mode of operation. In the heating mode, the heating coil valve is the active actuator and is modulated to maintain the supply air temperature at set point and the AHU dampers are controlled to allow the minimum outdoor air needed to satisfy the ventilation requirements. When the cooling load increases, the system simply mixes outdoor (cold) air and returns air to achieve the set point with both heating and cooling coil valves being closed. The mode of operation changes to mechanical cooling when the outdoor air is too warm to achieve the supply air temperature set points. In this mode, the cooling coil valve is manipulated to meet the supply air temperature set point. If the outdoor

air temperature is less than a certain value (typically 65 ^0F) the outdoor air damper is kept fully open. If the the outdoor air temperature is greater than the selected value, mechanical cooling is continued with the minimum outdoor air required for ventilation. The conditioned supply air is distributed to the five zones. Each zone has a variable-air-volume (VAV) box with hydronic reheat. In the next section, we describe the control structure in the VAV boxes.

VAV boxes

Figure 5.2 shows a schematic diagram of a zone VAV box and the corresponding sensors in the model. The VAV box uses a damper to modulate the amount of air entering the zone, and the hydronic coil to reheat the air entering the zone when necessary. The thermostat and flow sensor measure the air temperature in the zone and the flow rate of air into the zone. A discharge air temperature sensor measures the temperature of the air stream entering the zone (see Schein and House (2003) for more details on the control structure of VAV boxes).

The control structure for VAV boxes is based on two different control loops for cooling and heating, respectively. In the cooling mode, a cascade control loop is implemented. The outer loop has the zone temperature as the controlled variable and the set point for air flow rate to the room as the manipulated variable. In the inner loop, the damper is modulated to reach the desired set point for flow rate. In the heating mode, the air flow rate to the room model is kept constant at 1.6 m^3/s and 0.8 m^3/s for occupied and unoccupied periods, respectively, and the room temperature is maintained by modulating the reheat valve. Note that in practice the flow rates depend on the design loads and ventilation requirements. The flow rates and temperature set

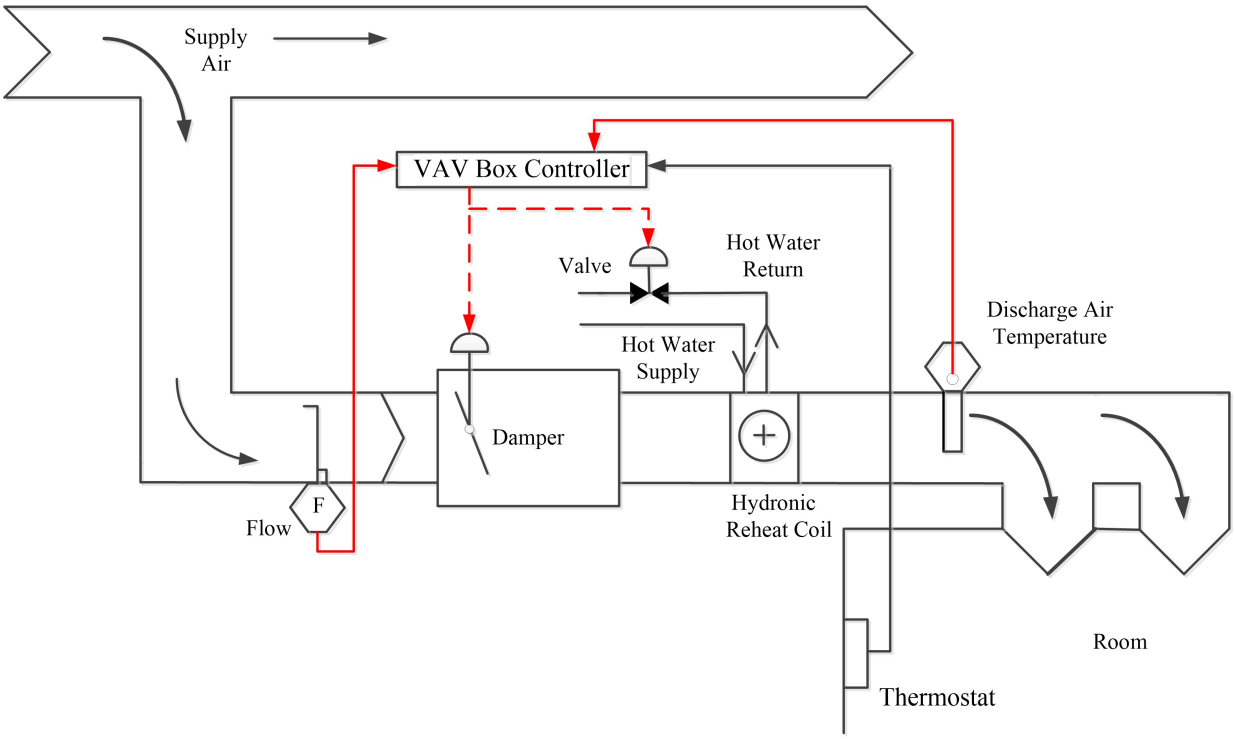


Figure 5.2: Schematic of a VAV box with hydronic reheat (recreated using the schematic from Schein and House (2003))

points used here are representative of the system considered. Switches between the heating and cooling mode are triggered to maintain the zone temperature between 21 ^0C (heating set point) and 23 ^0C (cooling set point). During the unoccupied period the dead band between the heating and cooling set point values is often widened in an effort to save energy. In our simulations, the unoccupied set points are 18 ^0C and 26 ^0C. As long as these conditions are satisfied, the controllers are inactive (i.e. the actuators remain in the same position), and any time the temperature goes beyond these values, the appropriate control action ensues.

The present work uses a detailed model for the HVAC systems as a test bed. In particular, the simulation platform modeled in Modelica is composed of at least 10 components and the interconnection of these components via mass and energy

balances. These models (and the resulting interconnections) are created using the in-built Modelica libraries, and results in a model with high dimensions (over 50 states). The ensuing FDI and safe-parking designs are developed using models identified via realistic measurements from the test bed (thus not all states are measured, and the measurements include measurement noise), and then implemented on the detailed simulation model.

There are several sources of faults in the VAV boxes including dampers or valves getting stuck, valve leakage and faulty sensor measurements. In this work, we focus on faults with severe effects on the performance of VAV boxes. Thus stuck dampers, stuck valves and biased sensor measurements are considered.

5.3 Fault diagnosis and fault handling design

There are several studies describing application of statistical based approaches for FDI in HVAC systems including single fault isolation in VAV boxes (Wang and Qin (2005), Qin and Wang (2005) and Du *et al.* (2007)). Our review did not find any papers where statistical or causal model based approaches have been used for isolation of multiple faults in VAV boxes of HVAC systems. In this work, we implement the combination of PCA and joint angle analysis methods as a basis for comparison with the proposed method. To this end, we first present a statistical model based FDI framework using a combination of PCA and joint angle analysis. Then we present a causal model based FDI scheme based on an identified linear time invariant (LTI) model and compare the two approaches. Note that statistical FDI design is the most common tool for diagnosing faults in HVAC systems. Finally, we illustrate a fault tolerant control design to handle stuck dampers in the zones of AHU model.

5.3.1 Statistical model based FDI design

In this section, we apply a combination of Hotelling's T^2 and SPE control charts for fault detection and a combination of contribution plots and joint angle analysis for fault isolation (see e.g., Yoon and MacGregor (2000), Yoon and MacGregor (2001) and Kourti (2005) for more information). The Hotelling's T^2 is used to detect variations in the plane of the first A principal components that is greater than what can be explained by the common cause variations or the so called outliers. In some sense, this metric evaluates the validity of the model for a particular observation. The SPE control chart, on the other hand, detects the new observations that can not be represented using the in-control model. Thus, if the SPE breaches the threshold, a fault is declared *only if the Hotelling's T^2 values is within the threshold.* (see e.g., Yoon and MacGregor (2000) and Kourti (2005)). Joint angle analysis is based on generating a fault library using fault signatures from the plant test or historical fault data and determining the measure of collinearity between the new measurements and the fault signatures (see e.g., Yoon and MacGregor (2001)). The cosine value between the new measurement vector and one of the known fault signatures gives the relative measure of collinearity between the two. Note that the joint angle analysis also utilizes PCA as it requires both principle component basis and residual basis to be determined (see e.g., Yoon and MacGregor (2001) for more on this). For convenience in the analysis, joint angle plots were introduced in Yoon and MacGregor (2000). The horizontal and vertical axises of the plot denote residual basis and model basis of the angle, respectively. A specific fault is declared if the measured angle goes to the top right corner (+1,+1) or bottom left corner (−1,−1). Note that as with most practical applications, noisy measurements must be filtered first to remove potentially random

variations (see e.g., Yoon and MacGregor (2001)).

Remark 5.1. Note that in this work PCA based statistical model based FDI techniques has been selected over partial least square (PLS) statistical model based FDI techniques owing to the fact that the PLS based statistical model based FDI techniques are only able to detect sensor faults (Negiz and Cinar (1997)) while PCA based techniques have the potential ability of sensor fault isolation.

Application of statistical model based FDI design

In this section, we apply the statistical model based approach described in Section 5.3.1 to the collected data from the detailed Modelica model for the VAV box of one of the five zones. To this end, at first we take the data of from the first two days of simulations with sampling time of one minute under healthy operation to build our in-control model.

The available data that we utilize includes measurements of effective valve opening (EVO), damper fractional opening (DVO), supply air temperature (SAT), supply fan pressure (SFP), air flow (AF) rate to the room and discharge air temperature (DAT), shown by red lines and red dashed lines in Figures 5.1 and 5.2. To reflect reality, the measurements are corrupted with white noise. The distribution of noises added to these measurements is normal with mean μ and standard deviation σ as listed in Table 5.1.

The noisy measurements are filtered using first order low pass filters before being fed to the controller. The filter parameters are described in Table 5.2. After normalizing the data, a principle component analysis is performed and using cross validation three principle components are found to be sufficient to represent 95% of the variance

Table 5.1: Noise distribution parameters

measurement	μ	σ	$unit$
SAT	0	0.08	^0C
SFP	0	2.5	pa
AF	0	0.2	m^3/s
DAT	0	0.2	^0C

Table 5.2: Filter parameters

measurement	Gain	Cut-off frequency (Hz)
SAT	1	0.0005
SFP	1	0.001
AF	1	0.001
DAT	1	0.001

in the data.

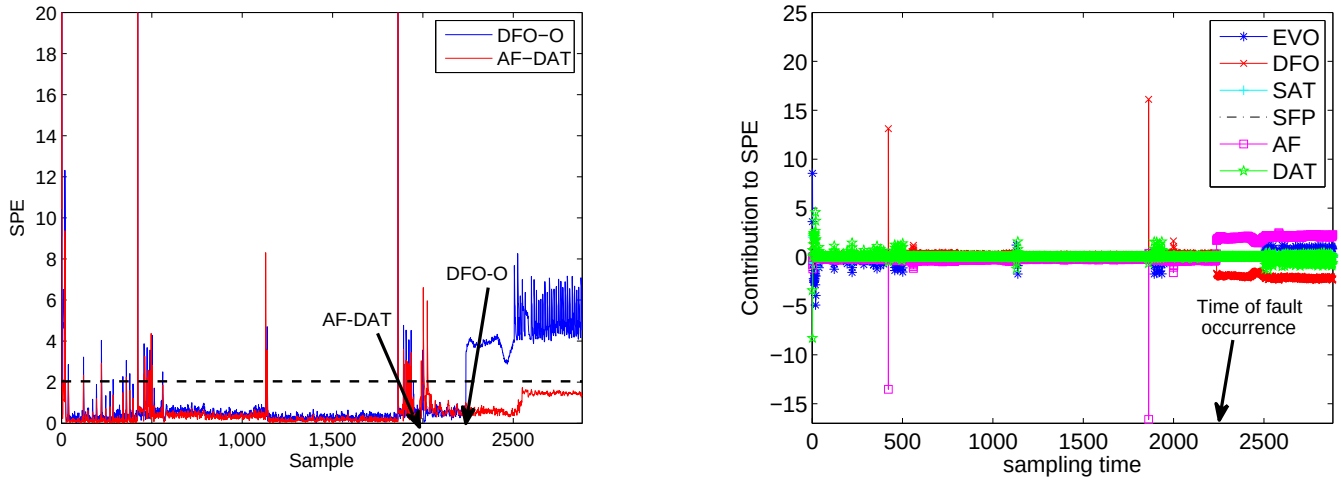
Next, we study the performance of the PCA based FDI design. To this end, we first consider the case where the damper gets stuck at 32% open position at 2233 sampling time (1:13 p.m. of second day of simulations). Note that both damper and valve positions range from 0 (fully closed) to 1 (fully open). Figure 5.3(a) shows the SPE control chart along with 99.5 % control limit. Note that the T^2 plot indicates existence of several outliers in the data with respect to the in-control model (the corresponding results are not presented here for sake of brevity). The SPE plot indicates the occurrence of some unusual events right after fault occurrence at 2234 sampling time ($1:14$ p.m) while a few false alarms are observed before fault occurrence due to the outliers identified in the T^2 plot. Thus the fault is successfully detected. The SPE contribution plot (Figure 5.3(b)) indicates both air flow (AF) and damper fractional opening (DFO) to be equally contributing to the event. The inability to isolate the fault is due to the model utilized in the the underlying FDI structure. For the

present fault scenario, a stuck damper causes the relationship between the prescribed (prescribed value refers to the control signal sent to the actuator) damper position and air flow to change. Thus both these are identified by the PCA based analysis as contributing to the unusual behavior, leading to the inability to isolate the fault. The same result was found when the damper was stuck at $DFO = 16\%$ position and is not presented here for sake of brevity.

We also consider another fault scenario where the valve gets stuck at the fully open position at 1931 sampling time (8:11 a.m. of day two of the simulation). It turns out that for the present case, the damper is able to be opened further to eliminate (or mask) the effect of this fault for some time until eventually the room temperature increases with respect to the healthy situation. Application of the PCA based analysis results in delayed detection of fault, but is not able to isolate the fault.

We next consider simultaneous positive bias faults in the flow sensor with magnitude of 2 and discharge air temperature sensor with magnitude of 6 taking place with at 1975 sampling time (8:55 a.m. of the day two of the simulation test). As can be seen from Figure 5.3, a fault is successfully detected while several false alarms are observed before fault occurrence due to outliers in the data. The SPE contribution plot indicates air flow (AF) sensor, damper fractional opening (DFO) and discharge air temperature (DAT) sensor as the variable contributing the most to the event, respectively, again indicating the inability to isolate the fault (the result is not presented here for sake of brevity). The method performs similarly for the case of multiple actuator faults.

Next, we apply joint angle analysis in an attempt to achieve fault isolation for the case studies that using a contribution plots was not successful. Note that the



(a) (b)

Figure 5.3: a) SPE plot for the cases when damper gets stuck at 32% position (DFO-O) and when air flow (AF) sensor and discharge air temperature (DAT) sensor are subject to simultaneous bias fault (AF-DAT). b) Contributions to SPE when the damper gets stuck at 32% position. The time of fault occurrence is indicated by arrows in both the figures.

application of joint angle analysis requires data subject to the occurrence of faults, with the faults being known. To this end, we generate a fault library using the available fault history data as listed in Table 5.3. The joint plots are generated from the time a fault is detected until five hours later.

For the first case described in Section 5.3.1 (the damper getting stuck at position 32%) the points generated by the pair of cosine values between the fault and the signatures of damper stuck at 32% position from the fault library are the dominating points at the top right corner. Thus, the fault is successfully isolated. Joint angle analysis also does not achieve fault isolation for the case of stuck damper at 16% position. Results are not presented here for sake of brevity.

We next consider the case when the valve gets stuck at fully open position as

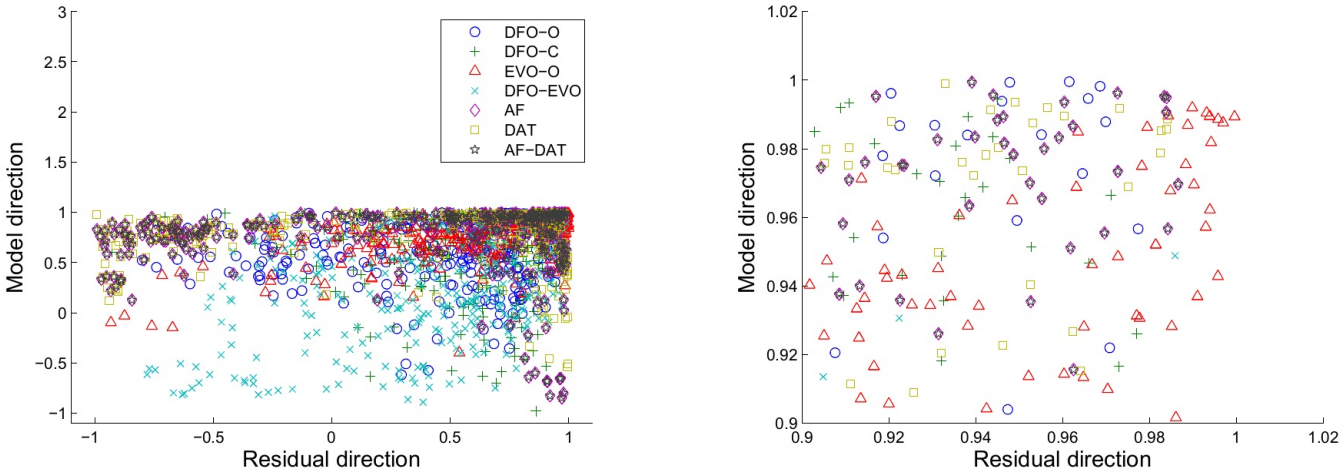
Table 5.3: Fault library

Fault	Abbreviation
Stuck damper 16 % open position	DFO-C
Stuck damper 32 % open position	DFO-O
Stuck valve at 100% open position	EVO-O
Simultaneous stuck damper at 16 % open position and stuck valve at 100% open position	DFO-EVO
Positive bias fault with magnitude of 2 m^3/s on flow	AF
Positive bias fault with magnitude of 6 ^0C on discharge air temperature	DAT
Simultaneous positive bias fault with magnitude of 2 m^3/s on air flow and positive bias fault with magnitude of 6 ^0C on discharge air temperature	AF-DAT

described in Section 5.3.1. As can be seen in Figure 5.4, the cosine values between the new fault and fault library signatures for the valve stuck fully open, bias fault on discharge air temperature sensor, simultaneous bias faults on flow and discharge air temperature sensors from the fault library are close to $(+1, +1)$. Thus, the joint angle analysis is not able to successfully isolate the fault.

For the case of bias faults occurring simultaneously on air flow (AF) and discharge air temperature (DAT) sensors (also described in Section 5.3.1) again, fault isolation is not achieved. The same results was found for simultaneous actuator faults (the results are not presented here for sake of brevity).

In summary, existing (non causal) PCA and joint angle analysis based approaches result in fault detection and isolation of some single faults, but are unable to isolate the faults when their effect is being masked by the control structure. As well, fault isolation is not achieved for the case of multiple actuator or multiple sensor faults.



(a) (b)

Figure 5.4: a: Joint plot using joint angle analysis for the case where valve gets stuck. b: Enlarged view of Figure 7(a)

5.3.2 Proposed model based FDI design

In this section, we design novel model based FDI filters for the VAV boxes. Intrinsic to the approach is the utilization of a causal dynamic model. To this end, first we utilize the subspace identification method to identify a LTI model for the VAV box. Next, we use the identified models to build the FDI filters.

Identifying a linear model for a zone

In this section, we identify a linear discrete time dynamic model of one of the zones (north zone) in the AHU simulation using the subspace identification method in Matlab (N4SID) (see Wang and Qin (2002) and Qin (2006) for more details on

this). To this end, a linear stochastic model of the following form is identified:

$$\begin{aligned} x(k+1) &= Ax(k) + Bu^*(k) + w(k) \\ y^*(k) &= Cx(k) + Du^*(k) \end{aligned} \tag{5.1}$$

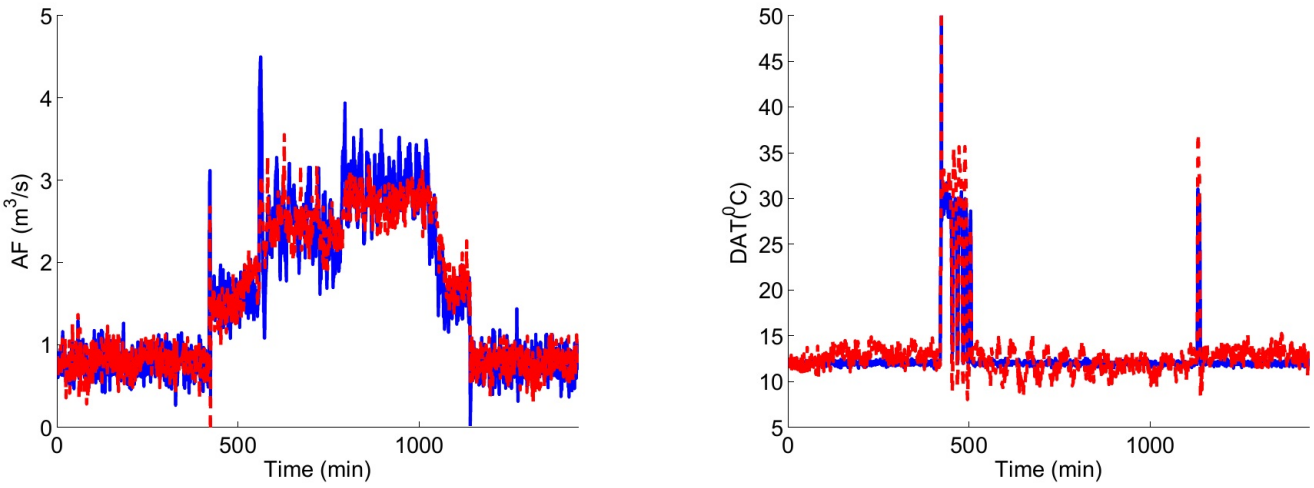
where $u^*(k)$, $y^*(k)$ and $x(k)$ are noise free inputs, noise free outputs and state variables and $w(k)$ denotes the process noise. In the identification approach, the available measurements for identification are assumed to be:

$$\begin{aligned} u(k) &= u^*(k) + o(k) \\ y(k) &= y^*(k) + v(k) \end{aligned} \tag{5.2}$$

where $o(k)$ and $v(k)$ denote the input and output white noise. Thus the task is to determine the order of the system and system matrices from past noisy input-output data.

The variables included in the identification were effective valve opening (EVO), damper fractional opening (DFO), supply air temperature (SAT) and supply fan pressure (SFP) as inputs and air flow (AF) to the room and discharge air temperature (DAT) to the zone as outputs.

We use data from the first day of the simulation test for identification and data from the second day of simulation test for validation. Note that in the training phase, the subspace identification approach determines initial values of the subspace states, and the system matrices. For a new/validation batch, these initial subspace states are unknown. Thus during this initial part (first 100 sample points for the present case), the identified model is used in conjunction with a Kalman filter, and takes the



(a) (b)

Figure 5.5: Model validation results using data of second day of simulations: Measured outputs (blue lines) and outputs generated by the identified model (red lines)

following from:

$$\begin{aligned} \hat{x}(k+1) &= A\hat{x}(k) + Bu(k) + Ke_k \\ \hat{y}(k) &= C\hat{x}(k) + Du(k) + e_k \end{aligned} \tag{5.3}$$

where $\hat{x}(k) \in \mathbb{R}^n$ denotes the vector of estimated subspace states, $u(k) \in \mathbb{R}^m$ denotes the vector of prescribed control inputs, $\hat{y}(k) \in \mathbb{R}^p$ denotes the vector of estimates of the output variables, K is the Kalman filter gain and $e_k = y(k) - C\hat{x}(k) - Du(k)$ is the estimation error.

The order of the identified model is picked to minimize the model identification error, resulting in a model with seven states. Figure 5.5 shows the model validation results. As can be seen from Figure 5.5, while there are errors (given that a linear model is being used to capture the dynamics of the detailed nonlinear model) the identified model captures zone dynamics reasonably well, making it a viable candidate for FDI design.

5.3.3 Fault detection and isolation design

Having identified a linear time invariant (LTI) model for the system, we next design the FDI filters. To this end, we first recognize that for the kinds of fault being considered, the evolution of the system subject to faults can be described by:

$$\begin{aligned} x(k+1) &= Ax(k) + B(u(k) + \tilde{u}(k)) + w(k) \\ y(k) &= Cx(k) + D(u(k) + \tilde{u}(k)) + \tilde{y}(k) + v(k) \end{aligned} \tag{5.4}$$

where $\tilde{u}(k) \in \mathbb{R}^m$ denotes the unknown fault vector for the actuators and $\tilde{y}(k)$ denotes the vector of unknown sensor faults. Due to the presence of physical constraints, the actual input implemented on the system, which is the sum of the (known) prescribed input and the (unmeasured) fault is also constrained. Let t_f denotes the time of fault occurrence and $\|\cdot\|$ the Euclidean norm for a vector. Note that under healthy operating conditions Eq. 5.3 is the equivalent innovation form of Eq. 5.4.

Actuator fault detection and isolation

To detect and isolate actuator faults, we estimate the implemented inputs by utilizing the measurements and the dynamic model, and compare the estimated value with the prescribed value. For the VAV box, we therefore design a filter that enables estimation of the effective valve opening and damper fractional opening using the results presented in Gillijns and De Moor (2007). By rearrangement, Eq. 5.4 becomes:

$$\begin{aligned} x(k+1) &= Ax(k) + B'u'(k) + G(d(k) + \tilde{d}(k)) + w(k) \\ y(k) &= Cx(k) + D'u'(k) + Hd(k) + v(k) \end{aligned} \tag{5.5}$$

where $u'(k) \in \mathbb{R}^{m_1}$ denotes the vector of known measurements as fault free inputs, $d(k) \in \mathbb{R}^{m_2}$ denotes the vector of prescribed inputs to be implemented to the plant by actuators where $m = m_1 + m_2$ and $\tilde{d}(k) \in \mathbb{R}^{m_2}$ denotes the vector of unknown actuator faults. In utilizing the filter, we need to employ the following assumption:

Assumption 5.1. Gillijns and De Moor (2007) The pair (A, C) is observable.

Assumption 5.2. Gillijns and De Moor (2007) Rank of H is m_2 i.e. $p \geq m_2$.

The filter consists of three steps as follows:

Estimation of unknown input:

$$\begin{aligned} \tilde{R}(k) &= CP^x_{k|k-1}C^T + R(k) \\ M(k) &= (H^T\tilde{R}^{-1}(k)H)^{-1}H^T\tilde{R}^{-1}(k) \\ \hat{d}(k) &= M(k)(y(k) - C\hat{x}_{k|k-1} - D'u'(k)) \\ P^d(k) &= (H^T\tilde{R}^{-1}(k)H)^{-1} \end{aligned} \tag{5.6}$$

Measurement update:

$$\begin{aligned} K(k) &= P^x_{k|k-1}C^T\tilde{R}^{-1}(k) \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + K(k)(y(k) - C\hat{x}_{k|k-1} - D'u'(k) - H\hat{d}(k)) \\ P^x_{k|k} &= P^x_{k|k-1} - K(k)(\tilde{R}_k - HP^d_kH^T)K^T(k) \\ P^{xd}_k &= (P^{dx}_k)^T = -K(k)HP^d \end{aligned} \tag{5.7}$$

Time update:

$$\begin{aligned}\hat{x}_{k+1|k} &= A\hat{x}_{k|k} + B'u'(k) + G_k\hat{d}(k)\\ P^x_{k+1|k} &= [A \quad G]\begin{bmatrix} P^x_{k|k} & P^{xd}_k \\ P^{dx}_k & P^d_k \end{bmatrix}\begin{bmatrix} A^T \\ G^T \end{bmatrix} + Q(k)\end{aligned} \tag{5.8}$$

where $Q(k) = E[w(k)w(k)^T] \geq 0$ and $R(k) = E[v(k)v^T(k)] > 0$.

The filter is initialized as follows:

$$\begin{aligned}\hat{x}_0 &= E[x_0]\\ P^x_0 &= E[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T]\end{aligned} \tag{5.9}$$

For $i = 1 \dots m_2$, the residuals are defined as below:

$$r_i(k) = \|d_i(k) - \hat{d}_i(k)\| \tag{5.10}$$

The FDI methodology using constant thresholds is presented in Theorem 1:

Theorem 5.1. *Consider the system of Eq. 5.5, for which Assumptions 5.1 and 5.2 hold. Then there exists δ_i such that if $r_i(k) > \delta_i$, then $\tilde{d}_{i,k} \neq 0$.*

Proof. Before fault occurrence i.e. $\tilde{d}_{i,k} = 0$, then

$$\begin{aligned}r_i(k) = \|d_i(k) - \hat{d}_i(k)\| &= \|(H^T\tilde{R}^{-1}(k)H)^{-1}H^T\tilde{R}^{-1}(k)(Cx_{k|k-1} - C\hat{x}_{k|k-1})\|\\ &= \|(H^T\tilde{R}^{-1}(k)H)^{-1}H^T\tilde{R}^{-1}(k)C\|\|\tilde{x}_{k|k-1}\|\end{aligned} \tag{5.11}$$

where $\tilde{x}_{k|k-1} = x_{k|k-1} - \hat{x}_{k|k-1}$ Since Assumptions 1 and 2 hold, according to Gillijns and De Moor (2007), the estimation error is bounded i.e $\|\tilde{x}_{k|k-1}\| \leq \delta_x$ where δ_x is a positive constant. Thus by selecting $\delta_i = \|(H^T\tilde{R}^{-1}(k)H)^{-1}H^T\tilde{R}^{-1}(k)C\|\delta_x$, if

Table 5.4: The residual notation, fault and the thresholds for the FDI design presented for actuator faults in Section 5.3.3 based on the framework in Section 5.3.3.

Residual	Faults	Threshold
r_1	$\tilde{u}_1$	0.644
r_2	$\tilde{u}_2$	0.182

$\tilde{d}_i(k) = 0$, then $r_i(k) \leq \delta_i$. Then, if $r_i(k) > \delta_i$, $\tilde{d}_i(k) \neq 0$. This concludes the proof of Theorem 5.1. □

Note that Theorem 5.1 addresses the problem of actuator fault detection and isolation for linear systems, and is utilized only as a guideline when implementing on the testbed, in particular in choosing the threshold. For the present example, to account for plant model mismatch, the thresholds are chosen as the maximum observed value for residuals after the input estimation filter has converged. In subsequent application for FDI $r_i(k) > \delta_i$ is considered a trigger only if this condition holds for at least ten consecutive samples.

Application of the actuators FDI framework

In this section, we apply the causal model based FDI method outlined in Section 5.3.3 to the VAV system. We first demonstrate the FDI capabilities assuming only actuator faults. Note that since neither effective valve opening (EVO) nor damper fractional opening (DFO) are subject to direct measurement noise (because they are values computed by the controller, and known), the values estimated by Eq. 5.6 are first filtered and then utilized for FDI. The filtering is carried out using moving average filters with span of 10 for both EVO and DFO. Table 5.4 shows the selected threshold corresponding to each residual as discussed in Section 5.3.3.

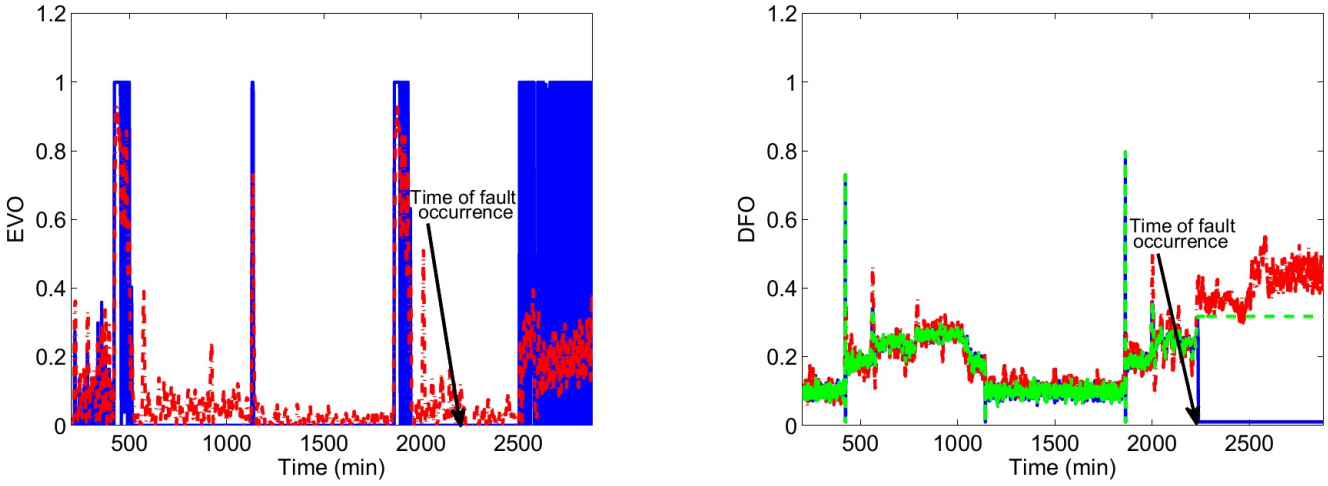
We again consider the damper fault scenario described in Section 5.3.1. The evolution of the actuators (valve and damper), their estimates provided by the estimation filter and the prescribed value are depicted by blue, red dashed, and green lines in Figure 5.6, respectively. As can be seen, the estimated values for actuators stay close to their actual values after fault occurrence. The evolution of residual profiles is shown in Figure 5.7. Since only r_2, the residual corresponding to a damper fault, breaches its threshold, the FDI filter successfully detects and isolates the fault. Note that although r_1 exceeds its threshold numerous times, it does so for fewer samples than needed to signal detection and isolation of a fault (each time for less than 10 consecutive samples). The same result holds for stuck damper at 0.16 position.

We next consider the case where the valve gets stuck at the fully open position as described in Section 5.3.1. As seen in Figures 5.8 and 5.9 only r_1, the residual corresponding to the faulty valve breaches its threshold leading to FDI. Recall that PCA and joint angle based approaches are unable to isolate this fault.

We next consider a case where both of the valve and damper get stuck simultaneously at 1874 sampling time (7:14 a.m. of second day of simulations) (also as in Section 5.3.1). The position for stuck valve and stuck damper are 1 and 0.16, respectively. The proposed approach results in FDI as seen in Figure 5.10.

Sensors fault detection and isolation

For sensor fault detection and isolation, we implement the methodology proposed in Du and Mhaskar (2014) considering a system with two outputs. For the identified linear system of Eq. 5.3, we design two Kalman filters each using only one of the sensors. To this end, let $y^i = C^i x(k) + D^i u(k) + v^i(k) + \tilde{y}^i(k) \in R$ where $i = 1, 2$

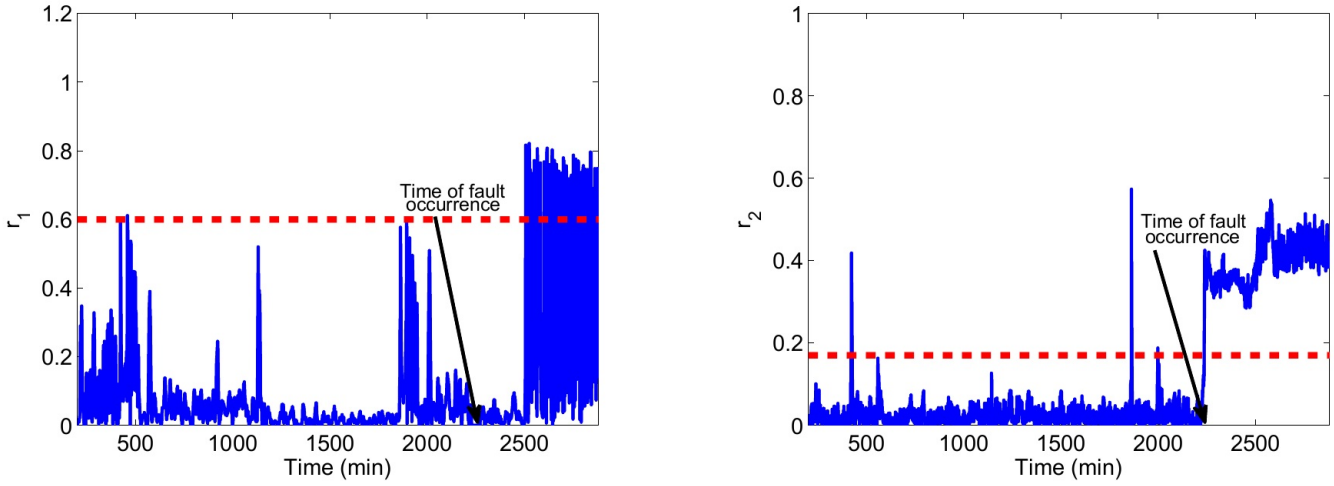


(a) (b)

Figure 5.6: Evolution of the prescribed value for actuators (solid blue lines), their estimates provided by the estimation filter (red dashed-dotted lines) and the actual value for actuators (green dashed lines). Note that the green dashed lines is plotted only when the actuator is stuck, and the actual value is different from the prescribed value. Otherwise only the blue line corresponding to prescribed value is plotted since the actual value implemented actuators is the same as prescribed value under healthy conditions. The damper gets stuck at 32% open position at 2233 sampling time (1:13 p.m. of second day of simulations) pointed by arrow in the both figures.

denotes the system output utilized by the ith Kalman filter, where $y^i = \bar{y}_i$, $C^i = \bar{C}_i$, $D^i = \bar{D}_i$, $v^i = \bar{v}_i$ and $\tilde{y}^i = \bar{\tilde{y}}_i$. To be able to design the ith Kalman filter using y^i, we need the following assumption:

Assumption 5.3. The pair (A, C^i) is observable.



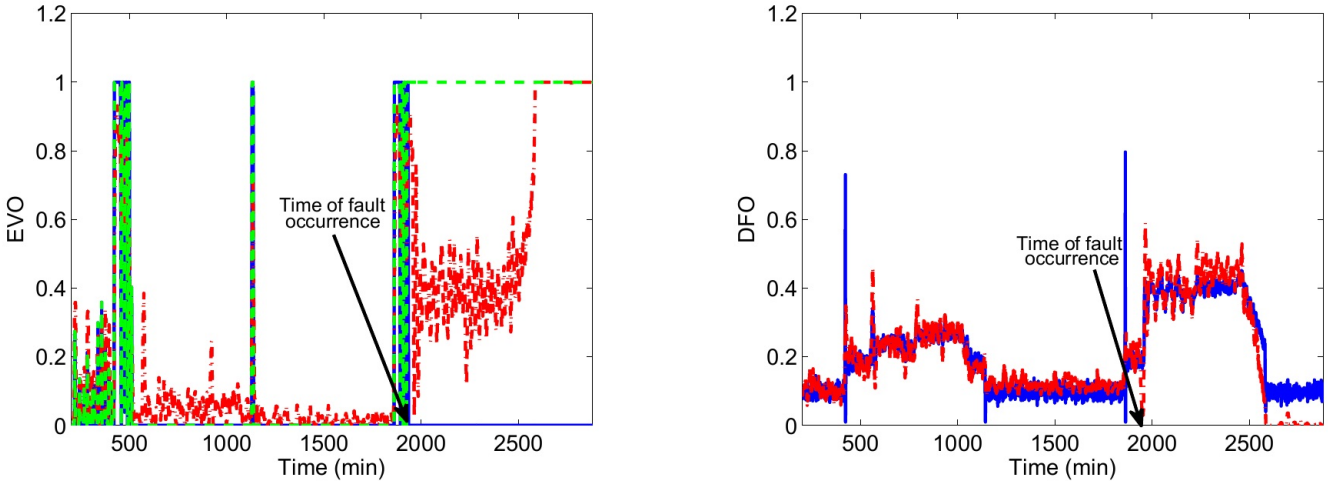
(a) (b)

Figure 5.7: Evolution of the residuals (blue solid lines) and thresholds (red dashed lines), when the damper gets stuck Top: Residual corresponding to fault at valve position. Bottom: Residual corresponding to fault at damper fractional opening. The damper gets stuck at 32% open position at 2233 sampling time (1:13 p.m. of second day of simulations) pointed by arrow in the both figures.

Thus the ith Kalman filter takes the following form:

$$\begin{aligned}
\tilde{R}_k^i &= C_i(AP_k^iA^T + Q^i)C_i^T + R_k^i \\
K^i &= (AP_k^iA^T + Q^i)C_i^T(\tilde{R}_k^i)^{-1} \\
\hat{x}^i(k) &= \hat{x}_{k|k-1} + K(y^i(k) - C_i\hat{x}_{k|k-1} - Du(k)) \\
P_{k+1}^i &= (I - K_iC_i)(AP_k^iA^T + Q^i) \\
\hat{x}_{k+1|k} &= A\hat{x}^i(k) + Bu(k)
\end{aligned} \tag{5.12}$$

where $Q(k) = E[w(k)w(k)^T] \geq 0$ and $R(k) = E[v(k)v^T(k)] > 0$. The filter is



(a) (b)

Figure 5.8: Evolution of the prescribed value for actuators (solid blue lines), their estimates provided by the estimation filter (red dashed-dotted lines) and the actual value for actuators (green dashed lines), if the corresponding actuator is subject to fault. Note that the green dashed lines is plotted only when the actuator is stuck, and the actual value is different from the prescribed value. In the absence of faults, the prescribed value equals the actual value and the lines overlap.The valve gets stuck at 100% open position at 1931 sampling time (8:11 a.m. of day two of the simulation) pointed by arrow in the both figures.

initialized as follows:

$$\begin{aligned} \hat{x}_0^i &= E[x_0^i] \\ P_0^i &= E[(x_0^i - \hat{x}_0^i)(x_0^i - \hat{x}_0^i)^T] \end{aligned} \tag{5.13}$$

Now, we describe the residual definition and fault detection and isolation design through the ith Kalman filter. Each residual is the norm of the difference between the expected trajectories of the state and the state estimates. The expected trajectories are calculated using the following prediction model:

$$\tilde{x}^i(k+1) = A\tilde{x}^i(k) + Bu(k) \quad t \in [t_{k-T}, t_k) \tag{5.14}$$

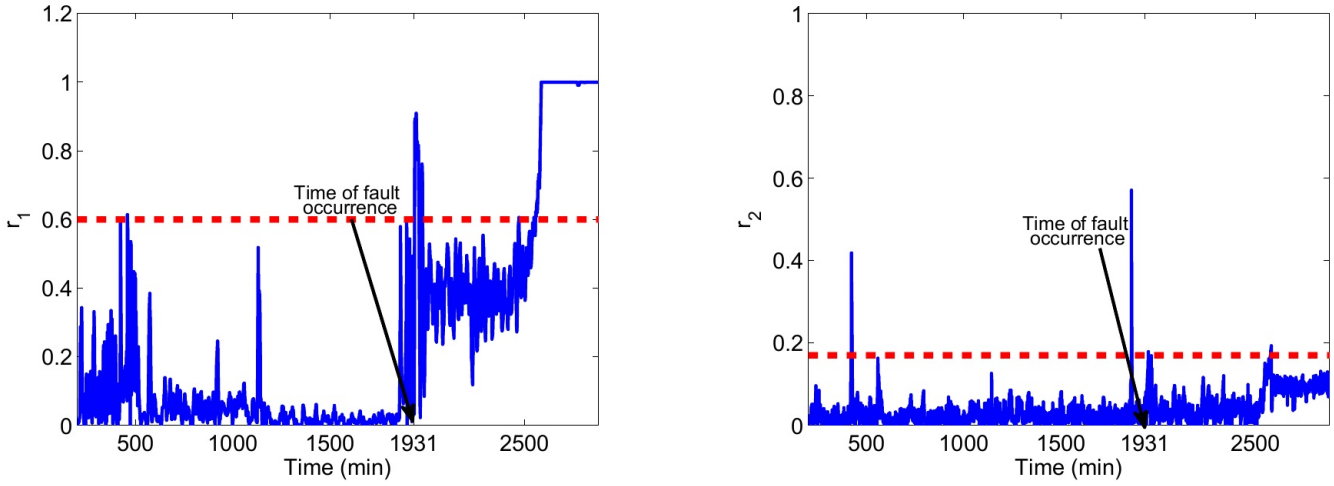
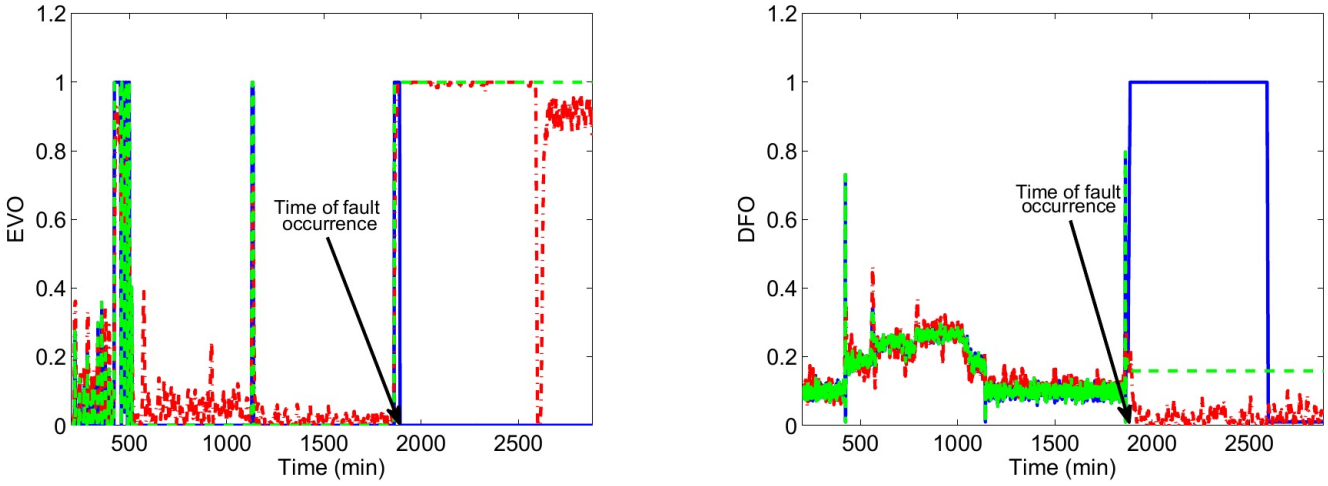


Figure 5.9: Evolution of the residuals (blue solid lines) and thresholds (red dashed lines). Top: Residual corresponding to fault at valve position. Bottom: Residual corresponding to fault at damper fractional opening. The valve gets stuck at 100% open position at 1931 sampling time (8:11 a.m. of day two of the simulation) pointed by arrow in the both figures.

where $\tilde{x}^i$ is the state of the prediction model, and T is the prediction horizon: $T = 1$ if $0 < t_k \leq t_{k'}$; $T = k - k'$ if $t_{k'} < t_k \leq t_{k'+T_p}$; and $T = T_p$ if $t_k > t_{k'+T_p}$, with a positive integer T_p being a chosen prediction horizon. The prediction model is initialized at the state estimate at time t_{k-T}: $\tilde{x}^i(k-T) = \hat{x}^i(k-T)$. By solving Eq. (5.14), the state prediction at time t_k is obtained. The corresponding residual is defined as below:

$$r_i(k) = \|\tilde{x}^i(k) - \hat{x}^i(k)\| \tag{5.15}$$

A fault is declared if at least one of the two residuals breach their thresholds. If only one of the residuals breaches its threshold, this means a fault has occurred in the corresponding sensor for the other residual (the insensitive residual) and the fault is successfully isolated. If both of the residuals breach their thresholds, this indicates



(a) (b)

Figure 5.10: Evolution of the prescribed value for actuators (solid blue lines), their estimates provided by the estimation filter (red dashed-dotted lines) and the true value for actuators (green dashed lines), if the corresponding actuator is subject to fault. The valve and damper get stuck simultaneously at 1874 sampling time (7:14 a.m. of day two of simulation) pointed by arrow in the both figures.

that both of the sensors are subject to fault. Thus, there is a unique breaching pattern corresponding to each fault scenario.

Remark 5.2. Note for a system with three or more outputs, breaching of all of the residuals only indicates detection of multiple faults and an additional number of observers must be designed for isolation of multiple faults in this case (see Du and Mhaskar (2014) and Du *et al.* (2013)).

Application of sensors FDI framework

In this section, we apply the described FDI scheme in Section 5.3.3 to the HVAC system of the north zone in the AHU model while considering only sensor faults. Since the model has two outputs, two Kalman filters are designed which use measurements

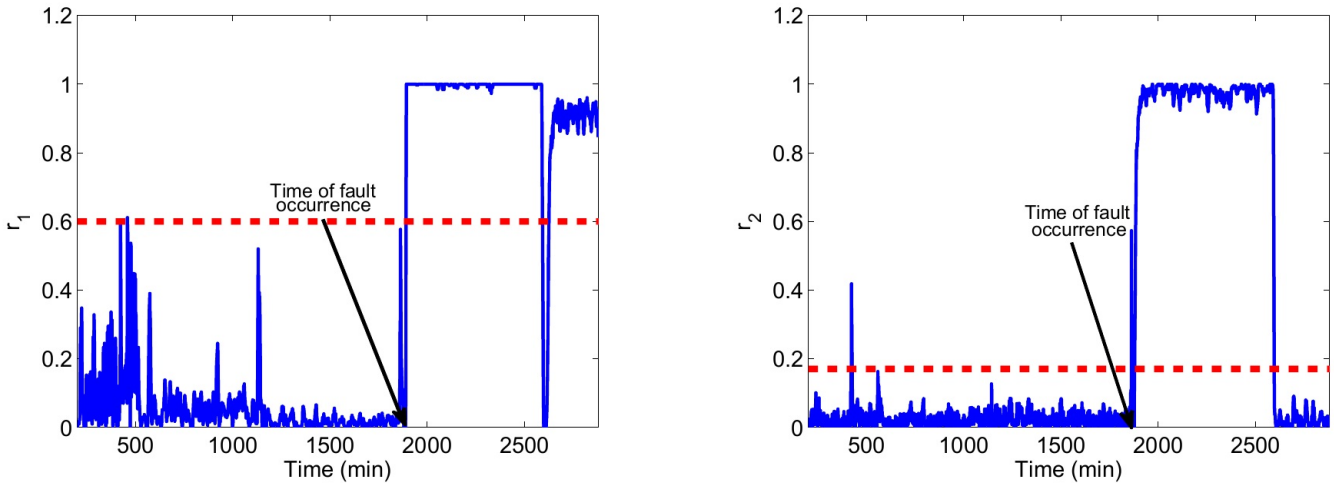


Figure 5.11: Evolution of the residuals (blue solid lines) and thresholds (red dashed lines). Top: Residual corresponding to fault at valve position. Bottom: Residual corresponding to fault at damper fractional opening. The valve and damper get stuck simultaneously at 1874 sampling time (7:14 a.m. of day two of testing) pointed by arrow in the both figures.

of discharge air temperature (DAT) and air flow (AF), respectively. To initialize the Kalman filters, for the first 100 sampling times each Kalman filter uses both measurements, then the additional measurement is removed. Also, for purpose of initialization the prediction model is associated with a Kalman filter that is active for 100 sampling times using both measurements. The Kalman filters parameters are $P_0 = O_{7\times7}$ and $Q = 10^3 I_{7\times7}$. The residuals are defined as described in 5.3.3. To account for effect of uncertainty, thresholds are selected as the maximum observed values for residuals while operating at steady state. Table 5.5 shows the selected threshold corresponding to each residual.

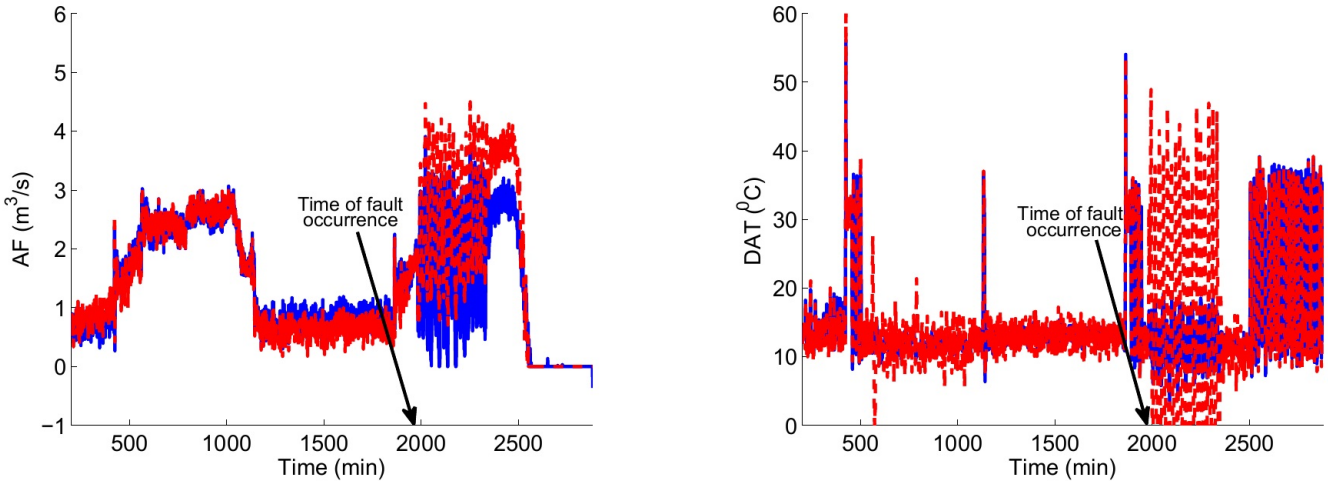
To this end, we consider simultaneous faults in the flow sensor and discharge air temperature sensor take place as described in Section 5.3.1. As can be seen from

Table 5.5: Faults to which the residuals are insensitive and thresholds for FDI design presented for sensor faults of the example in Section 5.3.3 based on the proposed framework in Section 5.3.3.

Residual	Faults	Threshold
r_1	$\tilde{y}_1$	0.155
r_2	$\tilde{y}_2$	0.23

Figure 5.12, since estimated values for both air flow and discharge air temperature are generated using faulty measurements, there is a discrepancy between the predicted values calculated using Eq. 5.14 and estimated values calculated using Eq. 5.12 for both flow and discharge air temperature. This causes both of the residuals to breach their thresholds (see Figure 5.13). Thus the fault is successfully detected and isolated. Note that the same results holds for single sensor faults. In conclusion, using the causal model based approaches enables isolation of multiple actuator and sensor faults (including those faults that are compensated by the controller). Note that for the case of simultaneous actuator and sensor faults, if the fault is detectable (see Remark 5.4), fault detection is still achieved. However, simultaneous actuator and sensor faults can not be isolated. This is due to fundamental limitation of the system dynamics and available measurements (see Remark 5.5).

Remark 5.3. Note that in this work a modified diagnostic observer methodology has been used for designing FDI scheme for diagnosing actuator faults. As mentioned earlier, the existing FDI approaches are not designed to detect and isolate actuator faults where the effect of the fault is compensated by the controller. The present design thus represents a modified diagnostic observer approach that uses the norm of the difference between the prescribed value and estimated value of inputs for residuals generation. In this way, the actuator faults can be diagnosed even in the presence of a



(a) (b)

Figure 5.12: Evolution of estimation (red) and prediction (blue) profiles for flow and discharge air temperature when both of the sensors are faulty. Simultaneous faults in the flow sensor and discharge air temperature sensor taking place at 1975 sampling time (8:55 a.m. of the day two of the simulation test) pointed by arrow in the both figures.

robust controller that compensates the effect of faults on the outputs. For sensor faults isolation, the diagnostic observer presented in Du and Mhaskar (2014) is utilized and is equivalent to other diagnostic approaches. The only difference is that the residuals are defined as norm of difference between expected and estimates trajectories of the plant.

Remark 5.4. Note since there are uncertainties associated with both of the identified linear model and the in-control model, a number of missed faults are inevitable by both of the proposed model based and statistical model based approaches utilized for FDI. This is due to the trade off between robustness and fault sensitivity (see e.g., Zhang *et al.* (2010b) and Dunia and Qin (1998) for more on detectability analysis for the proposed model based and statistical model based approaches, respectively). Note

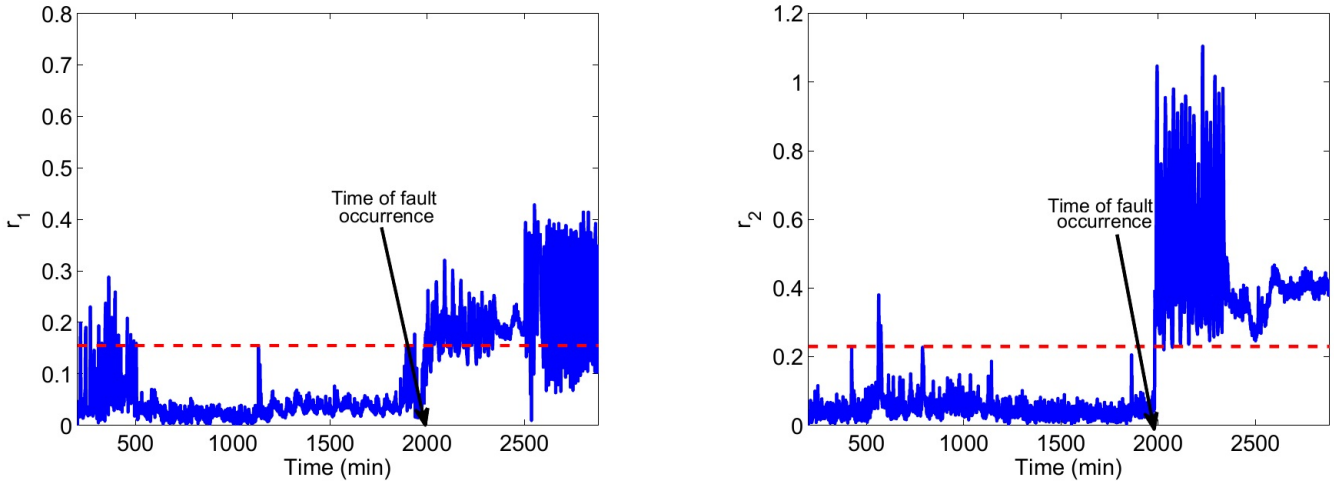


Figure 5.13: Evolution of the residuals (blue solid lines) and thresholds (red dashed lines). Top: Residual corresponding to fault at flow sensor. Bottom: Residual corresponding to fault at discharge air temperature. Simultaneous faults in the flow sensor and discharge air temperature sensor taking place at 1975 sampling time (8:55 a.m. of the day two of the simulation test) pointed by arrow in the both figures.

that a fault can be detected if and only if it satisfies the corresponding detectability conditions (for more on this, see Zhang *et al.* (2010b), Zhang (2011) and Dunia and Qin (1998)).

Remark 5.5. Simultaneous actuator and sensor faults can not be isolated for the present case, since the model structure does not meet the necessary and sufficient conditions for a fault scenario to be distinguishable (for more on this see Shahnazari *et al.* (2016)). Note that this is due to the fundamental limitation of the AHU system dynamics and limited measurements, not a limitation of the utilized FDI methodology.

Remark 5.6. The results obtained using the proposed model based approach advances the statistical model based approaches in terms of diagnosis of complex faults, multiple faults as well as applicability to the ability to isolate faults whose effect is

masked by the operating controller. Note that the present work does not consider zone temperature sensors in the FDI analysis, thus the zone temperature is not necessary in the identified statistical and causal models. Performing FDI for a larger section of the HVAC system remains the focus of future work.

5.3.4 Fault handling design

In this section, we illustrate a safe parking strategy for multi-zone HVAC systems to handle the fault of a stuck damper at a position close to zero but not completely closed (0.16 position). In this scenario, due to increasing cooling loads, the zone temperature increases and the control design can not handle the faulty situation. As a consequence the zone temperature exceeds the cooling set point during occupied periods. The safe parking strategy is based on replacing damper fractional opening with the supply fan pressure set point as the manipulated variable in the inner loop of the cascade control design for the faulty zone. By increasing the supply fan pressure set point, the air flow rate through the stuck damper will also increase. Thus, upon isolation, the faulty zone is parked in a temporary operating condition with new control objective (the cooling set point is now 24 ^0C rather than 23 ^0C) until the faulty equipment has been repaired or replaced. Figure 5.14 shows the evolution of air flow rate entering the north zone, damper fractional opening of a fault free zone, north zone temperature (ZT) profile and supply fan pressure in the presence and absence of the safe parking strategy. As shown in Figure 5.14, by increasing the supply fan pressure, the amount of air flow rate through the faulty damper increases. In the healthy zones, the dampers close more to maintain the same amount of air flow rate entering the healthy zones as before fault occurrence. In the faulty zone, since

the damper is stuck, more air flow enters the zone and as a result the temperature in the faulty zone is kept closer to the nominal operating condition until the system is brought back to its original condition. Note that there exist three possible scenarios for the case under consideration. One is to let the system operate in the faulty condition without an attempt to handle the fault, leading to increased discomfort, albeit without additional energy usage. The second scenario is to compensate for the fault completely by achieving the set point in the healthy condition (by increasing the static pressure in the entire system). In this case, the comfort is entirely retained at the price of more energy usage. The third scenario is to use the described safe parking strategy that trades off between the other two scenarios. Compared to the first scenario, there is improved comfort but more energy usage. Compared to the second scenario, there is decreased comfort, but also less energy usage. A more detailed implementation of the safe-parking approach to the HVAC system remains an objective of future work.

Remark 5.7. Note that when a VAV box damper gets stuck, different scenarios could take place depending on damper position and operating point of the systems. If the damper is stuck fully open, normally the VAV heating valve opens up and compensates for the fault effect at a price of additional energy usage in the system. In such a scenario, a possible safe-parking implementation would be to change the supply fan pressure (reduce it) but just enough so discomfort in the zone under consideration is reduced, without causing discomfort in the other zones. There is another scenario when the damper is stuck closed. In this case, nothing can be done until the faulty equipment is recovered or repaired. The particular case considered in the work is applicable for the situation where the damper gets stuck at an intermediate

position, and for the case where the controller is calling for the damper to open further. Note that the problem of fault handling in the context of VAV boxes is addressed, recognizing that while in the present context, the faults are not safety critical in nature, they do present an opportunity for trading off between comfort and energy usage. Thus, a safe parking strategy is designed to handle stuck dampers and the resulting energy reduction demonstrated.

5.4 Conclusion

In this work, we designed and implemented an integrated framework for fault diagnosis and fault handling in VAV boxes of HVAC systems. To this end, first, a statistical model based FDI scheme is designed using a combination of PCA and joint angle analysis. Then we designed a linear causal model based frameworks for detection and isolation of multiple actuator and multiple sensor faults. The linear model is identified using a subspace identification method. The causal model based frameworks achieve fault isolation for the cases that the statistical model based framework is not able to isolate the fault. Finally, we illustrated a FTC strategy to handle a stuck damper fault in one of the zones using a safe-parking approach to achieve a trade off between comfort and energy usage.

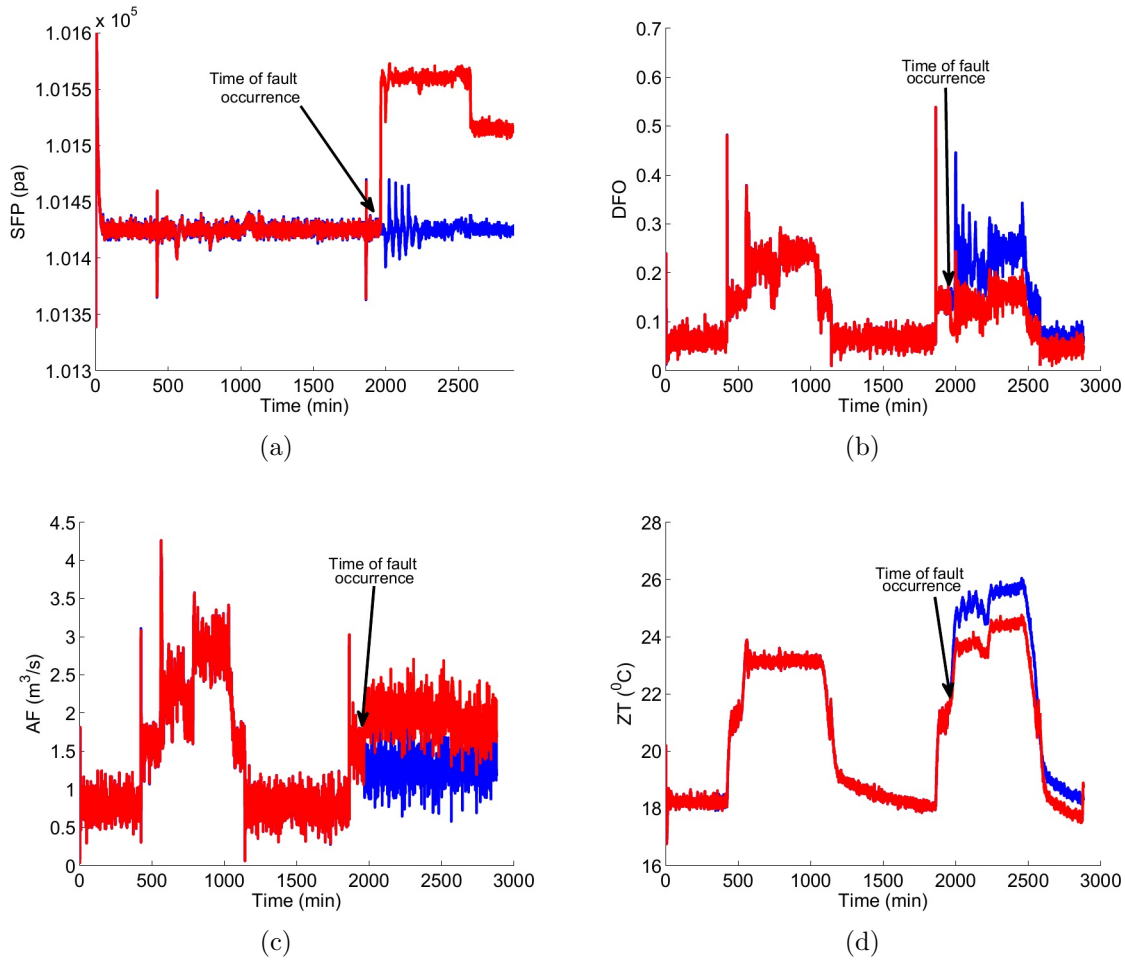


Figure 5.14: Evolution of a: supply fan pressure, b: damper fractional opening of the core zone, c: air flow rate entering the north zone , d: north zone temperature profile. Blue: When the safe parking framework is not active, Red: When the safe parking framework is active (time of fault occurrence is shown by arrow in the figures).

Chapter 6

Conclusion and future work

This chapter summarizes the main contributions of this thesis and suggests research opportunities for future work.

6.1 Conclusion

This work considers the problem of fault diagnosis and fault tolerant control for complex process systems. In Chapter 2, the problem of isolating distinguishable actuator and sensor faults in the solution copolymerization of MMA and VAc was considered. To achieve fault detection and isolation for the distinguishable faults in copolymerization reactor, an actuator and sensor fault detection and isolation framework was designed. To this end, first state estimates were generated using a bank of high-gain observers and then nonlinear fault detection and isolation (FDI) residuals were defined. The ability of the proposed framework in detecting and narrowing the possible locations for indistinguishable fault scenarios to a subset of possible scenarios was proved and verified through simulations. Illustrative linear FDI filters

were also designed for the purpose of comparison. While linear model based FDI only achieved fault detection, the application of the proposed FDI mechanism was found to also successfully isolate distinguishable faults even in the presence of plant-model mismatch and measurement noise.

In Chapter 3, the problem of actuator and sensor fault detection and isolation was addressed for control affine nonlinear systems subject to uncertainty. An FDI framework was proposed and fault detectability and isolability conditions were rigorously derived. Finally, the efficacy of the fault isolation framework subject to uncertainty and measurement noise was illustrated using a chemical reactor example.

In Chapter 4, the problem of simultaneous fault diagnosis was addressed for nonlinear uncertain networked systems utilizing a distributed fault detection and fault isolation strategy. The idea is to design a bank of local robust FDI schemes in a distributed manner with each FDI scheme corresponding to a subsystem. Time-varying thresholds were selected by explicitly accounting for the effect of uncertainties and faults in shared interconnections. In this way, robustness of the LFDI schemes to false alarms is guaranteed. Also, in the case of faults in the shared interconnections that can be isolated locally, the distributed architecture of the proposed FDI framework allows the other FDI schemes to function as intended. The detectability and isoability conditions were rigorously derived for the distributed FDI scheme. Effectiveness of the proposed methodology was shown via application to a reactor-separator process subject to uncertainty and measurement noise.

In Chapter 5, an integrated framework was designed and implemented for fault diagnosis and fault handling in VAV boxes of HVAC systems. To this end, first, a statistical model based FDI scheme is designed using a combination of PCA and joint

angle analysis. Then we designed linear causal model based frameworks for detection and isolation of multiple actuator and multiple sensor faults. The linear model is identified using a subspace identification method. The causal model based frameworks achieve fault isolation for the cases that the statistical model based framework is not able to isolate the fault. Finally, we illustrated a FTC strategy to handle a stuck damper fault in one of the zones using a safe-parking approach to achieve a trade off between comfort and energy usage.

6.2 Future work

The results of this thesis suggest the following topics for future work:

1. Actuator and sensor fault detection and isolation of nonlinear uncertain systems subject to delay

2. Actuator and sensor fault detection and isolation of nonlinear stochastic systems subject to uncertainty

3. Fault diagnosis and safe parking design for HVAC systems with supply fan pressure reset strategy

First, we consider the problem of fault diagnosis for nonlinear process systems subject delay. There are a few results available in the literature for FDI design in the presence of delay (see e.g., Chen and Saif (2006) and Yao *et al.* (2014)). However, none of these results have addressed the problem of fault detection and isolation in the presence of input, state and output delay. The other limitations of the existing results in the literature are considering only single actuator faults and being only

applicable to specific classes of nonlinear systems. To address this problem, the first step is to design a predictor that enables estimation of system states in the past and present in the presence of delay. The closed loop convergence property of the proposed predictor must be established rigorously. The next step is to design FDI filters in a way that they account for system complexities that cause nonzero values for the generated residuals in the absence of fault. The FDI filters must be designed in a way that enable differentiation between actuator and sensor faults.

Second, we consider the problem of fault diagnosis for nonlinear uncertain stochastic systems. There are only a few results available in the literature when it comes to FDI design for stochastic systems. In Keller (1999), FDI filters are designed to diagnose single and multiple actuator faults in linear stochastic systems using modified full order Kalman filters. In George (2012), a robust FDI scheme is designed for detection and isolation of single and simultaneous actuator and sensor faults in uncertain linear stochastic systems using robust observers. However, there is a lack of results in the literature when the problem of FDI design for uncertain nonlinear stochastic systems has been addressed. To investigate this problem, the first step is to design a robust observer with bounded estimation error in the presence of uncertainty and process noise. Again, the closed loop convergence property of the proposed estimator must be rigorously established. The next step is to design FDI filters in a way they explicitly account for system complexities that cause nonzero values in the generated residuals in the absence of fault. As far as defining residuals is concerned, the methodology presented in Du *et al.* (2013) can be adapted accordingly. The third step is to establish the ability of the proposed FDI scheme in detecting and isolating faults in the presence of uncertainty rigorously that can be achieved via proper

detectability and isoability analysis, respectively.

Third, we consider the problem of fault diagnosis and safe parking design for HVAC systems with supply fan pressure reset strategy. Static pressure setpoint reset strategy is based on reseting the static pressure setpoint for the zone requiring the most pressure i.e., the setpoint is decreased until one of the zone dampers is open at a desired amount. The main advantage of utilizing the reset strategy is energy saving since the static pressure is being kept at the minimum value in the allowable range and it increases only when it is required (see e.g., Taylor (2007)). However, when the supply fan pressure is not constant, the relation between damper fractional opening (DFO) and air flow (AF) becomes nonlinear. Also, when one of the dampers gets stuck at almost closed positions, the VAV controller signals the damper to be fully open as long as long the damper is stuck. Thus the supply fan pressure reset strategy does not continue to work properly in this case and there is some energy loss in the system due to increase in supply fan pressure values. As a result of this, a fault handling strategy is required to minimize both discomfort and energy losses in the zones. To address these problems, the first step is to identify a model that can capture the nonlinear behavior of the system. As model identification is concerned, an online model identification strategy (see e.g., Alanqar *et al.* (2017)) will be used that results in a linear parameter varying (LPV) model for the system. The second step is to design a FDI framework. The methodologies proposed in Du *et al.* (2013) or Shahnazari *et al.* (2018) can be utilized for FDI design depending on the structure of the identified model for the system. The third step is to design a fault handling framework for stuck dampers with ability to balance between providing comfort in the zones and energy usage.

Appendix A

Appendix

Appendix A: Lyapunov Based Model Predictive Control

Consider the nonlinear system of Eq. 2.4 with input constraints for which a control Lyapunov function V exists. Let Π denote a set of states where $\dot{V}(x(t))$ can be made negative by using the allowable values of the constrained input:

$$\Pi = \{x \in \mathbb{R}^n : L_f V(x) + \inf_u (L_G V(x) u) \leq -\varepsilon^{**}\} \tag{A.1}$$

where $L_G V(x) = [L_{g_1} V(x), \ldots, L_{g_m} V(x)]$, with g_i the ith column of G and ε^{**} is a positive real number. The controller of Mahmood and Mhaskar (2008) possesses a stability region, an estimate of which is given by

$$\Omega = \{x \in \Pi : V(x) \leq c_{max}\}, \tag{A.2}$$

where c_{max} is a positive (preferably the largest possible) constant. Having defined the sets Π, Ω, the Lyapunov based predictive controller of Mahmood and Mhaskar (2008) follows the formulation below:

$$u_{MPC}(x) := argmin\{J(x,t,u(.))|u(.) \in S\} \tag{A.3}$$

$$s.t. \quad \dot{x} = f(x) + G(x)u \tag{A.4}$$

$$\dot{V} \leq -\varepsilon^*, \forall \tau \in [t, t+\Delta), \quad if \quad V(x(t)) > \delta' \tag{A.5}$$

$$\dot{V} \leq -\delta', \quad \forall \tau \in [t, t+\Delta), \quad if \quad V(x(t)) \leq \delta' \tag{A.6}$$

$$x(t+\tau) \in \Pi, \quad \forall \tau \in [t, t+\Delta) \quad if \quad V > c_{max} \tag{A.7}$$

where $S = S(t,T)$ is the family of piecewise continuous functions (functions continuous from the right), with T denoting the control horizon, mapping $[t, t+T]$ into U, for a given positive number d, δ' is a positive real number such that $V(x) \leq \delta'$ implies that $\|x(t)\| \leq d$ and ϵ^* is a positive real number (related to ε^{**} through Δ, see Mahmood and Mhaskar (2008) for details). A control $u(.)$ in S is characterized by the sequence $\{u(t_k)\}$ and satisfies $u(\tau) = u(t_k)$ for all $\tau \in [t_k, t_k + \Delta]$. The objective function is given by

$$J(x,t,u(.)) = \int_t^{t+T} [\|x^u(s;x,t)\|^2_{Q_w} + \|u(s)\|^2_{R_w}]ds \tag{A.8}$$

where Q_w and R_w are positive semidefinite, and strictly positive definite, symmetric matrices, respectively, $x^u(s;x,t)$ denotes the solution of Eq. A.4, due to control u, with initial state x at time t and T is specified horizon. In accordance with the

receding horizon implementation, the minimizing control u_{MPC} is then applied to the system over $[t, t + \Delta]$, and the same procedure is repeated at the next instant. The stability property of the Lyapunov based predictive control design in Mahmood and Mhaskar (2008) can be formulated as follows: given any positive real number d, there exists a positive real number Δ^* such that if $\Delta \in [0, \Delta^*]$ and $x(0) \in \Omega$ then $x(t) \in \Omega$, for all $t \geq 0$ and $\lim sup_{t\to\infty}\|x(t)\| \leq d$. Furthermore, for $x(0) \in \Pi\backslash\Omega$, if the optimization problem of Eqs A.3-A.7 is successively feasible, then $x(t) \in \Pi$ $\forall t \geq 0$ and $\lim sup_{t\to\infty}\|x(t)\| \leq d$ (see e.g. Mahmood and Mhaskar (2008) for further details). Note that the control design in Mahmood and Mhaskar (2008) is used only to illustrate the proposed framework in this work and the obtained results hold under any control law that guarantees stability of the closed loop system.

Bibliography

Afram, A. and Janabi-Sharifi, F. (2014). Review of modeling methods for HVAC systems. *Applied Thermal Engineering*, **67**(1), 507–519.

Ahrens, J. H. and Khalil, H. K. (2009). High-gain observers in the presence of measurement noise: A switched-gain approach. *Automatica*, **45**(4), 936–943.

Alanqar, A., Durand, H., and Christofides, P. D. (2015). On identification of well-conditioned nonlinear systems: Application to economic model predictive control of nonlinear processes. *AIChE Journal*, **61**(10), 3353–3373.

Alanqar, A., Durand, H., and Christofides, P. D. (2017). Error-triggered on-line model identification for model-based feedback control. *AIChE Journal*, **63**(3), 949–966.

Armaou, A. and Demetriou, M. A. (2008). Robust detection and accommodation of incipient component and actuator faults in nonlinear distributed processes. *AIChE journal*, **54**(10), 2651–2662.

Atassi, A. N. and Khalil, H. K. (1999). A separation principle for the stabilization of a class of nonlinear systems. *Automatic Control, IEEE Transactions on*, **44**(9), 1672–1687.

Bengea, S. C., Li, P., Sarkar, S., Vichik, S., Adetola, V., Kang, K., Lovett, T., Leonardi, F., and Kelman, A. D. (2015). Fault-tolerant optimal control of a building HVAC system. *Science and Technology for the Built Environment*, **21**(6), 734–751.

Benosman, M. and Lum, K.-Y. (2010). Passive actuators' fault-tolerant control for affine nonlinear systems. *IEEE Transactions on Control Systems Technology*, **18**(1), 152–163.

Blanke, M., Kinnaert, M., Lunze, J., Staroswiecki, M., and Schröder, J. (2006). *Diagnosis and fault-tolerant control*, volume 691. Springer.

Boem, F., Ferrari, R. M., Keliris, C., Parisini, T., and Polycarpou, M. M. (2017). A distributed networked approach for fault detection of large-scale systems. *IEEE Transactions on Automatic Control*, **62**(1), 18–33.

Chen, J. and Patton, R. J. (2012). *Robust model-based fault diagnosis for dynamic systems*, volume 3. Springer Science & Business Media.

Chen, W. and Saif, M. (2006). An iterative learning observer for fault detection and accommodation in nonlinear time-delay systems. *International Journal of Robust and Nonlinear Control*, **16**(1), 1–19.

Chen, Y. and Lan, L. (2010). Fault detection, diagnosis and data recovery for a real building heating/cooling billing system. *Energy Conversion and Management*, **51**(5), 1015–1024.

Chilin, D., Liu, J., Muñoz de la Peña, D., Christofides, P. D., and Davis, J. F. (2010). Detection, isolation and handling of actuator faults in distributed model predictive control systems. *Journal of Process Control*, **20**(9), 1059–1075.

Clark, R. (1978a). Instrument fault detection. *IEEE Transactions on Aerospace Electronic Systems*, **14**, 456–465.

Clark, R. N. (1978b). A simplified instrument failure detection scheme. *Aerospace and Electronic Systems, IEEE Transactions on*, (4), 558–563.

Cole, W. J., Hale, E. T., and Edgar, T. F. (2013). Building energy model reduction for model predictive control using openstudio. In *American Control Conference (ACC), 2013*, pages 449–454. IEEE.

Cole, W. J., Powell, K. M., Hale, E. T., and Edgar, T. F. (2014). Reduced-order residential home modeling for model predictive control. *Energy and Buildings*, **74**, 69–77.

Congalidis, J. P., Richards, J. R., and Ray, W. H. (1989). Feedforward and feedback control of a solution copolymerization reactor. *AIChE Journal*, **35**(6), 891–907.

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

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

Doymaz, F., Chen, J., Romagnoli, J. A., and Palazoglu, A. (2001a). A robust strategy for real-time process monitoring. *Journal of Process Control*, **11**(4), 343–359.

Doymaz, F., Romagnoli, J. A., and Palazoglu, A. (2001b). A strategy for detection and isolation of sensor failures and process upsets. *Chemometrics and Intelligent Laboratory Systems*, **55**(1), 109–123.

Du, M. (2012). *Fault diagnosis and fault-tolerant control of chemical process systems.* Ph.D. thesis, McMaster University.

Du, M. and Mhaskar, P. (2011). A safe-parking and safe-switching framework for fault-tolerant control of switched nonlinear systems. *International Journal of Control*, **84**(1), 9–23.

Du, M. and Mhaskar, P. (2013). Active fault isolation of nonlinear process systems. *AIChE Journal*, **59**(7), 2435–2453.

Du, M. and Mhaskar, P. (2014). Isolation and handling of sensor faults in nonlinear systems. *Automatica*, **50**(4), 1066–1074.

Du, M., Gandhi, R., and Mhaskar, P. (2011). An integrated fault detection and isolation and safe-parking framework for networked process systems. *Industrial & Engineering Chemistry Research*, **50**(9), 5667–5679.

Du, M., Nease, J., and Mhaskar, P. (2012). An integrated fault diagnosis and safe-parking framework for fault-tolerant control of nonlinear systems. *International Journal of Robust and Nonlinear Control*, **22**(1), 105–122.

Du, M., Scott, J., and Mhaskar, P. (2013). Actuator and sensor fault isolation of nonlinear process systems. *Chemical Engineering Science*, **104**, 294–303.

Du, Z. and Jin, X. (2007). Detection and diagnosis for multiple faults in VAV systems. *Energy and Buildings*, **39**(8), 923–934.

Du, Z., Jin, X., and Wu, L. (2007). Fault detection and diagnosis based on improved PCA with JAA method in VAV systems. *Building and Environment*, **42**(9), 3221–3232.

Dunia, R. and Qin, S. J. (1998). Joint diagnosis of process and sensor faults using principal component analysis. *Control Engineering Practice*, **6**(4), 457–469.

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

Fekih, A. (2014). Fault-tolerant flight control design for effective and reliable aircraft systems. *Journal of Control and Decision*, **1**(4), 299–316.

Ferrari, R. M., Parisini, T., and Polycarpou, M. M. (2012). Distributed fault detection and isolation of large-scale discrete-time nonlinear systems: An adaptive approximation approach. *IEEE Transactions on Automatic Control*, **57**(2), 275–290.

Findeisen, R., Imsland, L., Allgöwer, F., and Foss, B. A. (2003). Output feedback stabilization of constrained systems with nonlinear predictive control. *International Journal of robust and nonlinear control*, **13**(3-4), 211–227.

Floquet, T., Barbot, J.-P., Perruquetti, W., and Djemai, M. (2004). On the robust fault detection via a sliding mode disturbance observer. *International Journal of control*, **77**(7), 622–629.

Frank, P. M. (1987). Fault diagnosis in dynamic systems via state estimation-a survey. In *System fault diagnostics, reliability and related knowledge-based approaches*, pages 35–98. Springer.

Frank, P. M. (1990). Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy: A survey and some new results. *Automatica*, **26**(3), 459–474.

Freeman, R. and Kokotovic, P. V. (2008). *Robust nonlinear control design: state-space and Lyapunov techniques*. Springer Science & Business Media.

Gandhi, R. and Mhaskar, P. (2008). Safe-parking of nonlinear process systems. *Computers & Chemical Engineering*, **32**(9), 2113–2122.

Gandhi, R. and Mhaskar, P. (2009). A safe-parking framework for plant-wide fault-tolerant control. *Chemical Engineering Science*, **64**(13), 3060–3071.

George, J. (2012). Robust fault detection and isolation in stochastic systems. *International Journal of Control*, **85**(7), 779–799.

Gertler, J. J. and Monajemy, R. (1995). Generating directional residuals with dynamic parity relations. *Automatica*, **31**(4), 627–635.

Gillijns, S. and De Moor, B. (2007). Unbiased minimum-variance input and state estimation for linear discrete-time systems. *Automatica*, **43**(1), 111–116.

Hao, X., Zhang, G., and Chen, Y. (2005). Fault-tolerant control and data recovery in HVAC monitoring system. *Energy and buildings*, **37**(2), 175–180.

House, J. M., Vaezi-Nejad, H., and Whitcomb, J. M. (2001). An expert rule set for fault detection in air-handling units. *Transactions-American Society of Heating Refrigerating and Air Conditioning Engineers*, **107**(1), 858–874.

Hu, Y. and El-Farra, N. H. (2011). Robust fault detection and monitoring of hybrid process systems with uncertain mode transitions. *AIChE Journal*, **57**(10), 2783–2794.

Isermann, R. (2006). *Fault-diagnosis systems: an introduction from fault detection to fault tolerance*. Springer Science & Business Media.

Kaboré, P., Othman, S., McKenna, T., and Hammouri, H. (2000). Observer-based fault diagnosis for a class of non-linear systems application to a free radical copolymerization reaction. *International Journal of Control*, **73**(9), 787–803.

Kabore, R. and Wang, H. (2001). Design of fault diagnosis filters and fault-tolerant control for a class of nonlinear systems. *IEEE transactions on Automatic Control*, **46**(11), 1805–1810.

Keliris, C., Polycarpou, M. M., and Parisini, T. (2015). Distributed fault diagnosis for process and sensor faults in a class of interconnected input–output nonlinear discrete-time systems. *International Journal of Control*, **88**(8), 1472–1489.

Keller, J.-Y. (1999). Fault isolation filter design for linear stochastic systems. *Automatica*, **35**(10), 1701–1706.

Khalil, H. K. (1999). High-gain observers in nonlinear feedback control. In *New directions in nonlinear observer design*, pages 249–268. Springer.

Khalil, H. K. (2002). Nonlinear systems. 3rd ed. *Prentice Hall, Upper Saddle River, NJ*.

Khalili, M. (2017). *Distributed Adaptive Fault-Tolerant Control of Nonlinear Uncertain Multi-Agent Systems*. Ph.D. thesis, Wright State University.

Kourti, T. (2005). Application of latent variable methods to process control and multivariate statistical process control in industry. *International Journal of adaptive control and signal processing*, **19**(4), 213–246.

Ljung, L. (1998). System identification. In *Signal analysis and prediction*, pages 163–173. Springer.

Ma, H.-J. and Yang, G.-H. (2012). Detection and adaptive accommodation for actuator faults of a class of non-linear systems. *IET Control Theory & Applications*, **6**(14), 2292–2307.

Ma, H.-J. and Yang, G.-H. (2013). Residual generation for fault detection and isolation in a class of uncertain nonlinear systems. *International Journal of Control*, **86**(2), 263–275.

Ma, J., Qin, J., Salsbury, T., and Xu, P. (2012). Demand reduction in building energy systems based on economic model predictive control. *Chemical Engineering Science*, **67**(1), 92–100.

Magni, J.-F. and Mouyon, P. (1994). On residual generation by observer and parity space approaches. *IEEE Transactions on Automatic Control*, **39**(2), 441–447.

Mahadevan, S. and Shah, S. L. (2009). Fault detection and diagnosis in process data using one-class support vector machines. *Journal of process control*, **19**(10), 1627–1639.

Mahmood, M. and Mhaskar, P. (2008). Enhanced stability regions for model predictive control of nonlinear process systems. *AIChE journal*, **54**(6), 1487–1498.

Mahmood, M., Gandhi, R., and Mhaskar, P. (2008a). Safe-parking of nonlinear process systems: Handling uncertainty and unavailability of measurements. *Chemical Engineering Science*, **63**(22), 5434–5446.

Mahmood, M., Gandhi, R., and Mhaskar, P. (2008b). Safe-parking of nonlinear process systems: Handling uncertainty and unavailability of measurements. *Chem. Eng. Sci.*, **63**, 5434 – 5446.

Mattei, M., Paviglianiti, G., and Scordamaglia, V. (2005). Nonlinear observers with H_∞ performance for sensor fault detection and isolation: a linear matrix inequality design procedure. *Contr. Eng. Prac.*, **13**, 1271–1281.

Mendoza-Serrano, D. I. and Chmielewski, D. J. (2012). HVAC control using infinite-horizon economic MPC. In *Decision and control (CDC), 2012 IEEE 51st annual conference on*, pages 6963–6968. IEEE.

Mendoza-Serrano, D. I. and Chmielewski, D. J. (2014). Smart grid coordination in building HVAC systems: EMPC and the impact of forecasting. *Journal of Process Control*, **24**(8), 1301–1310.

Mhaskar, P. (2006). Robust model predictive control design for fault-tolerant control of process systems. *Industrial & engineering chemistry research*, **45**(25), 8565–8574.

Mhaskar, P., Gani, A., McFall, C., Christofides, P. D., and Davis, J. F. (2007). Fault-tolerant control of nonlinear process systems subject to sensor faults. *AIChE Journal*, **53**(3), 654–668.

Mhaskar, P., McFall, C., Gani, A., Christofides, P. D., and Davis, J. F. (2008). Isolation and handling of actuator faults in nonlinear systems. *Automatica*, **44**(1), 53–62.

Negiz, A. and Cinar, A. (1997). Monitoring sensor performance in multivariable

continuous processes. In *American Control Conference, 1997. Proceedings of the 1997*, volume 1, pages 314–318. IEEE.

Nimmo, I. (1995). Adequately address abnormal operations. *Chemical engineering progress*, **91**(9), 36–45.

Peng, D., El-Farra, N. H., Geng, Z., and Zhu, Q. (2015). Distributed data-based fault identification and accommodation in networked process systems. *Chemical Engineering Science*, **136**, 88–105.

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

Qin, J. and Wang, S. (2005). A fault detection and diagnosis strategy of VAV air-conditioning systems for improved energy and control performances. *Energy and buildings*, **37**(10), 1035–1048.

Qin, S. J. (2006). An overview of subspace identification. *Computers & chemical engineering*, **30**(10), 1502–1513.

Reppa, V., Polycarpou, M. M., and Panayiotou, C. G. (2015). Distributed sensor fault diagnosis for a network of interconnected cyberphysical systems. *IEEE Transactions on Control of Network Systems*, **2**(1), 11–23.

Sánchez-Peña, R. S., Casín, J. Q., and Cayuela, V. P. (2007). *Identification and control: the gap between theory and practice*. Springer Science & Business Media.

Schein, J. and House, J. M. (2003). Application of control charts for detecting faults in

variable-air-volume boxes. *Transactions-American society of heating, refrigerating and air-conditioning engineers*, **109**(2), 671–682.

Schein, J., Bushby, S. T., Castro, N. S., and House, J. M. (2006). A rule-based fault detection method for air handling units. *Energy and Buildings*, **38**(12), 1485–1492.

Schön, T. B., Wills, A., and Ninness, B. (2011). System identification of nonlinear state-space models. *Automatica*, **47**(1), 39–49.

Schoukens, M. and Tiels, K. (2017). Identification of block-oriented nonlinear systems starting from linear approximations: A survey. *Automatica*, **85**, 272–292.

Seem, J. E. (2001). Integrated control and fault detection of HVAC equipment. US Patent 6,223,544.

Shahnazari, H. and Mhaskar, P. (2016). Simultaneous actuator and sensor fault isolation of nonlinear systems subject to uncertainty. In *American Control Conference (ACC), 2016*, pages 6857–6862. IEEE.

Shahnazari, H., Mhaskar, P., *et al.* (2016). Fault detection and isolation analysis and design for solution copolymerization of MMA and VAc process. *AIChE Journal*, **62**(4), 1054–1064.

Shahnazari, H., Mhaskar, P., House, J. M., and Salsbury, T. I. (2018). Heating, ventilation and air conditioning systems: Fault detection and isolation and safe parking. *Computers & Chemical Engineering*, **108**, 139 – 151.

Talukdar, A. and Patra, A. (2010). Dynamic model-based fault tolerant control of variable air volume air conditioning system. *HVAC&R Research*, **16**(2), 233–254.

Tan, C. P. and Edwards, C. (2003). Sliding mode observers for robust detection and reconstruction of actuator and sensor faults. *International Journal of Robust and Nonlinear Control*, **13**(5), 443–463.

Tayarani-Bathaie, S. S., Vanini, Z. S., and Khorasani, K. (2014). Dynamic neural network-based fault diagnosis of gas turbine engines. *Neurocomputing*, **125**, 153–165.

Taylor, S. T. (2007). Increasing efficiency with VAV system static pressure setpoint reset. *ASHRAE Journal*, **49**(6), 24.

Tong, C., El-Farra, N. H., Palazoglu, A., and Yan, X. (2014). Fault detection and isolation in hybrid process systems using a combined data-driven and observer-design methodology. *AIChE Journal*, **60**(8), 2805–2814.

Touretzky, C. R. and Baldea, M. (2014a). Integrating scheduling and control for economic mpc of buildings with energy storage. *Journal of Process Control*, **24**(8), 1292–1300.

Touretzky, C. R. and Baldea, M. (2014b). Nonlinear model reduction and model predictive control of residential buildings with energy recovery. *Journal of Process Control*, **24**(6), 723–739.

Van Overschee, P. and De Moor, B. (2012). *Subspace identification for linear systems: Theory-Implementation-Applications*. Springer Science & Business Media.

Vemuri, A. T. and Polycarpou, M. M. (1997). Robust nonlinear fault diagnosis in input-output systems. *International Journal of Control*, **68**(2), 343–360.

Venkatasubramanian, V., Rengaswamy, R., Yin, K., and Kavuri, S. N. (2003). A review of process fault detection and diagnosis: Part i: Quantitative model-based methods. *Computers & chemical engineering*, **27**(3), 293–311.

Wang, J. and Qin, S. J. (2002). A new subspace identification approach based on principal component analysis. *Journal of process control*, **12**(8), 841–855.

Wang, S. and Qin, J. (2005). Sensor fault detection and validation of VAV terminals in air conditioning systems. *Energy Conversion and Management*, **46**(15), 2482–2500.

Wu, S. and Sun, J.-Q. (2011). Cross-level fault detection and diagnosis of building HVAC systems. *Building and Environment*, **46**(8), 1558–1566.

Yan, X.-G. and Edwards, C. (2007). Nonlinear robust fault reconstruction and estimation using a sliding mode observer. *Automatica*, **43**(9), 1605–1614.

Yan, X.-G. and Edwards, C. (2008). Robust decentralized actuator fault detection and estimation for large-scale systems using a sliding mode observer. *International Journal of control*, **81**(4), 591–606.

Yang, J. and Zhu, F. (2015). FDI design for uncertain nonlinear systems with both actuator and sensor faults. *Asian Journal of Control*, **17**(1), 213–224.

Yang, J., Zhu, F., Wang, X., and Bu, X. (2015). Robust sliding-mode observer-based sensor fault estimation, actuator fault detection and isolation for uncertain nonlinear systems. *International Journal of Control, Automation and Systems*, **13**(5), 1037–1046.

Yao, L., Wang, H., and Cocquempot, V. (2014). Adaptive nonlinear actuator fault

diagnosis for uncertain nonlinear systems with time-varying delays. *Asian Journal of Control*, **16**(4), 1057–1065.

Yin, X. and Liu, J. (2017). Distributed output-feedback fault detection and isolation of cascade process networks. *AIChE Journal*.

Yoon, S. and MacGregor, J. F. (2000). Statistical and causal model-based approaches to fault detection and isolation. *AIChE Journal*, **46**(9), 1813–1824.

Yoon, S. and MacGregor, J. F. (2001). Fault diagnosis with multivariate statistical models part i: using steady state fault signatures. *Journal of process control*, **11**(4), 387–400.

Yoshida, H., Kumar, S., and Morita, Y. (2001). Online fault detection and diagnosis in VAV air handling unit by RARX modeling. *Energy and Buildings*, **33**(4), 391–401.

Zhang, X. (2011). Sensor bias fault detection and isolation in a class of nonlinear uncertain systems using adaptive estimation. *Automatic Control, IEEE Transactions on*, **56**(5), 1220–1226.

Zhang, X. and Zhang, Q. (2012). Distributed fault diagnosis in a class of interconnected nonlinear uncertain systems. *International Journal of Control*, **85**(11), 1644–1662.

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

Zhang, X., Parisini, T., and Polycarpou, M. M. (2005). Sensor bias fault isolation

in a class of nonlinear systems. *Automatic Control, IEEE Transactions on*, **50**(3), 370–376.

Zhang, X., Polycarpou, M. M., and Parisini, T. (2010a). Adaptive fault diagnosis and fault-tolerant control of MIMO nonlinear uncertain systems. *International Journal of Control*, **83**(5), 1054–1080.

Zhang, X., Polycarpou, M. M., and Parisini, T. (2010b). Fault diagnosis of a class of nonlinear uncertain systems with lipschitz nonlinearities using adaptive estimation. *Automatica*, **46**(2), 290–299.

Zhang, X., Zhang, Q., and Sonti, N. (2011). Diagnosis of process faults and sensor faults in a class of nonlinear uncertain systems. *Systems Engineering and Electronics, Journal of*, **22**(1), 22–32.

Zhu, F. and Yang, J. (2013). Fault detection and isolation design for uncertain nonlinear systems based on full-order, reduced-order and high-order high-gain sliding-mode observers. *International Journal of Control*, **86**(10), 1800–1812.