DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

A FILTERING APPROACH FOR FAULT DIAGNOSIS OF NONLINEAR UNCERTAIN SYSTEMS

CHRISTODOULOS KELIRIS

A Dissertation Submitted to the University of Cyprus in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

February, 2015

Christodoulos Keliris

Christodoulos Keliris

© Christodoulos Keliris, 2015

VALIDATION PAGE

Doctoral Candidate: Christodoulos Keliris

Doctoral Thesis Title: A Filtering Approach for Fault Diagnosis of Nonlinear Uncertain Systems

The present Doctoral Dissertation was submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy at the ***Department of Electrical and Computer Engineering****, and was approved on* ***10.02.2015*** *by the members of the* ***Examination Committee****.*

Examination Committee:

Committee Chair: _____
Christos Panayiotou, Associate Professor

Research Supervisor: _____
Marios M. Polycarpou, Professor

Committee Member: _____
Christoforos Hadjicostis, Professor

Committee Member: _____
Constantinos Pattichis, Professor

Committee Member: _____
Thomas Parisini, Professor

Christodoulos Keliris

DECLARATION OF DOCTORAL CANDIDATE

The present doctoral dissertation was submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy of the University of Cyprus. It is a product of original work of my own, unless otherwise mentioned through references, notes, or any other statements.

Doctoral Candidate: Christodoulos Keliris

Signature: .....

Christodoulos Keliris

Περίληψη

Αυτή η διατριβή προτείνει μια καινοτόμα μέθοδο διάγνωσης σφαλμάτων, μέσω γραμμικών φίλτρων, για μη γραμμικά συστήματα με μη ακριβή μοντελοποίηση και κάτω από την ύπαρξη θορύβου στις μετρήσεις. Στις περισσότερες πρακτικές εφαρμογές, η ύπαρξη αβεβαιότητας στη μοντελοποίηση και θορύβου στις μετρήσεις μπορεί να επηρεάσει σημαντικά την απόδοση των μεθόδων ανίχνευσης σφαλμάτων είτε μέσω λανθασμένων συναγερμών είτε μέσω αποτυχημένων ανιχνεύσεων. Ως εκ τούτου, η επίτευξη μιας ισορροπίας μεταξύ της ευρωστίας και της ορθής ανιχνευσιμότητας σε μια μέθοδο ανίχνευσης σφαλμάτων είναι εξαιρετικά σημαντική. Όπως διαφαίνεται από αυτή τη διατριβή, η προτεινόμενη μέθοδος ανίχνευσης σφαλμάτων προσφέρει σημαντικά οφέλη και προς τις δύο αυτές κατευθύνσεις, εγγυώντας από τη μία τη μη ύπαρξη λανθασμένων συναγερμών και διατηρώντας ταυτόχρονα τα όρια ανίχνευσης σφαλμάτων σε χαμηλά επίπεδα. Αυτό επιτυγχάνεται μέσω της ενσωμάτωσης στο σχεδιασμό των σημάτων υπολοίπων (***residuals***) και ορίων (***thresholds***) μιας γενικής κλάσης γραμμικών φίλτρων, που βοηθούν στην εξασθένηση και μετριασμό του θορύβου. Οι ιδιότητες της προτεινόμενης μεθόδου ερευνώνται κάτω από ένα αυστηρά μαθηματικό υπόβαθρο και οδηγούν σε συμπεράσματα σχετικά με το μέγεθος των ανιχνεύσιμων σφαλμάτων, το μέγιστο χρόνο ανίχνευσης ενός σφάλματος και προσδιορίζοντας τη σχέση ανάμεσα στο χρόνο ανίχνευσης και την τάξη και τους πόλους των φίλτρων που χρησιμοποιούνται. Η προτεινόμενη μεθοδολογία ανίχνευσης σφαλμάτων αναπτύσσεται για όλους τους συνδυασμούς συνεχούς/διακριτού χρόνου και της περίπτωσης ύπαρξης όλων μετρήσεων των μεταβλητών κατάστασης/περίπτωσης εισόδου-εξόδου. Επιπλέον, η προτεινόμενη μέθοδος παρουσιάζει σημαντικές ιδιότητες σχετικά με τη μεταφορά των επιδράσεων των σφαλμάτων μεταξύ των αλληλοσυνδεόμενων υποσυστημάτων σε κατανεμημένα συστήματα, κάτι που επιτρέπει την εξαγωγή χρήσιμων συμπερασμάτων σχετικά με τα σφάλματα που έχουν παρουσιαστεί, επιτρέποντας έτσι την εκπόνηση μιας υψηλού επιπέδου μεθόδου απομόνωσης των σφαλμάτων. Τέλος, χρησιμοποιούνται προσαρμοστικές μέθοδοι προσέγγισης για την εκμάθηση της αβεβαιότητας στη μοντελοποίηση, έτσι

Christoforos Kallis

ώστε να επιτευχθούν καλύτερα όρια και να ενισχυθεί η ανιχνευσιμότητα των σφαλμάτων. Η τεχνική εκμάθησης ενοποιείται με τη χρήση των φίλτρων και ως εκ τούτου προκύπτει ένα ενοποιημένο πλαίσιο που εκμεταλλεύεται τα οφέλη που προκύπτουν τόσο από την εκμάθηση της αβεβαιότητας όσο και από τη χρήση των φίλτρων για εξασθένηση του θορύβου. Η χρήση των τεχνικών εκμάθησης χρησιμοποιείται στις περιπτώσεις που υπάρχουν σφάλματα στη διεργασία του συστήματος (***process faults***) ή στις μετρήσεις (***sensor faults***), επιτρέποντας τον καθορισμό του τύπου του σφάλματος που έχει παρουσιαστεί ενώ παρέχεται επίσης και εκτίμησή του. Σε όλες τις περιπτώσεις που περιλαμβάνονται σε αυτή τη διατριβή, παρουσιάζονται οι αντίστοιχες συνθήκες ανιχνευσιμότητας σφαλμάτων, και διαφαίνεται η αποτελεσματικότητα της προτεινόμενης μεθόδου ανίχνευσης σφαλμάτων μέσω αποτελεσμάτων προσομοίωσης.

Christodoulos Kellis

Abstract

This dissertation proposes a novel filtering approach for the problem of fault diagnosis for nonlinear systems with modeling uncertainty and measurement noise. In most real world applications the presence of modeling uncertainty and measurement noise may influence significantly the performance of fault detection schemes by causing, either missed fault detections or false alarms and therefore, this trade-off between robustness and detectability is of crucial importance. The proposed filtering approach for fault diagnosis offers beneficial characteristics in both directions by guaranteeing no false alarms while at the same time providing tight detection thresholds. This is achieved by embedding into the design of the residual and threshold signals a general class of filters which takes advantage of the filtering noise suppression properties. The properties of this novel approach are rigorously investigated, providing results regarding the magnitude of the detected faults, an upper bound on the detection time and the relation of the detection time with respect to the order and pole locations of the filters used. The proposed fault detection filtering approach is devised under all the combinations of continuous/discrete time and the full-state measurement and input-output case. Furthermore, the proposed distributed fault diagnosis approach encompasses important characteristics regarding fault propagation among interconnected subsystems which allows the derivation of a high-level fault isolation scheme. Finally, the devised filtering framework is exploited to integrate adaptive approximation methods for learning the modeling uncertainty in addition to the use of filtering for measurement noise attenuation, in order to obtain tighter thresholds and enhance fault detection. The use of learning is used in the cases of process and sensor faults allowing the identification of the fault type and estimation of the fault. In all the cases considered in this dissertation, respective fault detectability conditions characterizing the class of detectable faults are obtained and, simulation results illustrate the effectiveness of the proposed fault filtering approach.

Christodoulos Keliris

Acknowledgments

I would like to express my sincerest gratitude to my advisor Prof. Marios M. Polycarpou for his exceptional guidance, support and availability. His advice and patience were significant elements that characterized the fruitful cooperation over the years. I would also like to thank Prof. Thomas Parisini for his valuable comments and feedback and also, the other members of the PhD Committee for their time in evaluating the work contained in this thesis. Finally, my warmest thanks go to my family for their love and support over the years.

Christodoulos Keliris

Contents

Christodoulos Keliris

Christodoulos Keliris

Christodoulos Keliris

List of Figures

Christodoulos Keliris

Christodoulos Kelinis

Chapter 1

Introduction

Fault diagnosis has been in the front end of the technological evolution for a few decades. As systems become more and more complex the need for enhanced robustness, fault tolerance and sustainability becomes of essential importance. Before we proceed it would be better to define what a fault is. A fault is anything that prevents the system directly or indirectly from conducting the job it was designed for. The direct way is by the malfunction of one or several system components which lead to the collapse of the overall process and this may be complete or partial as in the case of degraded performance. The indirect way is again by malfunction of system components but this time they seem to be working but in fact they provide false information. For instance, a simple example would be a sensor giving wrong measurements. These sort of faults are the most difficult to detect and also they are the most dangerous ones as they can remain hidden until a major breakdown occurs. Hence, the smooth and reliable operation of large scale systems is of critical importance in the modern technological world. Potential failures could lead to major catastrophes and consequently could trigger a chain of failing dependent systems such as electric power systems, communication and water networks, along with production plants causing a tremendous economic and social damage. Therefore, the safe and reliable operation of such systems through the early detection of any "small" fault before they become serious failures is a crucial component of the overall system performance and sustainability.

Generally, an important issue that is often overseen is the presence of measurement noise and modeling uncertainty. In most real world applications such uncertainties may influence significantly the performance of fault detection schemes by causing, for example, false alarms. On the other hand, simply acknowledging the presence of these uncertainties and by using constant or functional bounds to deal with them, imposes the risk in leading to conser-

Christodoulos Keliris

vative fault detection schemes. In these cases, the designed detection thresholds can be too conservative and as a result, missed faults can occur. Therefore, there is the need to design robust fault detection schemes to counteract the effects of the measurement noise and the presence of the modeling uncertainty, so that tighter detection thresholds can be obtained, whilst guaranteeing no false alarms.

Most approaches for fault detection and accommodation so far have been based on a centralized architecture, in which information about the state of the system are gathered and processed centrally. Motivated by advances in communications and distributed sensing, there has recently been significant interest in distributed and hierarchical fault diagnosis methods. In many cases, a distributed FDI framework is not an option but a necessity, since many factors contribute to this formulation such as the large scale nature of the system to be monitored, its spatial distribution, the inability to access certain parts of the system from a remote monitoring component and therefore local diagnosis should be performed. The distributed nature of the underlying systems' process poses significant issues regarding fault isolation once a fault is detected. Essentially, the first line of defense of any distributed fault diagnosis scheme once a fault is detected, is the identification of the faulty subsystem.

In addition, the faults that can occur in a system can be both process and sensor ones. The research community usually focuses only on one fault type (process or sensor) and proposes schemes specifically targeted to that type, something that encompasses significant risks for the system reliability since the other fault type is ignored. Therefore, designing robust fault diagnosis schemes able to deal concurrently with both types of faults is a crucial component for system reliability and sustainability. Moreover, a useful characteristic of such schemes would be the identification of the fault type and the generation of an estimation of the fault so that it can be used subsequently for fault accommodation purposes.

1.1 Motivation and objectives of the present work

Motivated by the issues raised above, this thesis proposes a robust fault detection scheme utilizing filtering, under a deterministic framework, for nonlinear systems with modeling uncertainty and noisy measurements. As a result, the fault detectability is enhanced whilst no false alarms are guaranteed. More specifically, the main contribution of this thesis is the development of a fault detection scheme utilizing a general class of filters which is embedded into the design of the residual and threshold signals in a way that takes advantage of the filtering noise suppression properties. Hence, the proposed approach relies on filtering

Christodoulos Kellinis

certain signals so that the effect of the measurement noise is attenuated, allowing for the design of less conservative thresholds. This filtering approach for fault detection in nonlinear systems is rigorously investigated in Chapter 3, providing results regarding the magnitude of detectable faults and the filtering impact (according to the poles' location and filters' order) on the detection time.

The devised filtering scheme is developed for a wide range of cases comprising of all the combinations of continuous/discrete time and considering the full state measurements/input-output case (Chapters 3-6). In all cases, respective fault detectability conditions are derived and simulation results illustrate the effectiveness of the proposed scheme. In these cases, in which interconnected, nonlinear systems are considered, the scheme is comprised of a set of interacting fault detection agents, in which each subsystem is monitored by its respective detection agent. The devised scheme is inherently distributed, since the fault detection agents monitoring their respective subsystems, exchange the measurements of the interconnection variables. This particular design, utilizing the measurements in the design of the local estimation models, is further investigated under the discrete-time, input-output case (Chapter 6) in which the consideration of sensor faults is also considered apart from the occurrence of process faults (Chapters 3-5 deal only with process faults). As it is shown in Chapter 6, the developed distributed fault diagnosis scheme possesses important isolation characteristics in the case considering both process and sensor faults, that aid in the extrapolation of further information regarding the type and location of the fault that occurred. Specifically, the approach deals with process and sensor faults in an integrated way and encompasses important characteristics regarding fault propagation among subsystems. As it is shown, a process fault that occurs in a subsystem can only be detected by its respective detection agent, whilst a sensor fault that occurs in a subsystem may also be detected by the detection agents of neighboring subsystems that are affected by the subsystem where the sensor fault occurred. This differentiating element is exploited in order to derive a high-level fault isolation scheme, able to provide information regarding the the type and location of the fault that has occurred. Therefore, the proposed distributed fault detection approach encompasses significant benefits for the fault isolation task which can be exploited by a more sophisticated isolation scheme to pinpoint the exact fault that occurred.

Finally, Chapter 7 develops an integrated filtering and adaptive approximation based approach for fault diagnosis of process and sensor faults in a class of continuous-time nonlinear systems with modeling uncertainties and measurement noise. The primary objective of this chapter, is the design of a unified fault diagnosis approach with: a) enhanced fault

Christodoulos Kellis

detectability characteristics through the integration of learning and filtering techniques and, b) fault identification capabilities that allow the identification of the fault type (process or sensor) and the estimation of the fault. To focus on the new challenges, only a single nonlinear system is considered (the interconnected subsystems can be considered as one large scale system). The first task of enhanced fault detectability, is achieved through the integration of learning and filtering techniques which allow the derivation of even tighter detection thresholds whilst maintaining the guarantee of no false alarms. This is accomplished by using adaptive approximation methods for learning the modeling uncertainty and, at the same time, filtering is used to dampen the effect of measurement noise. Both filtering and learning are integrated in a unified framework and intertwined through the devised filtering approach, which is decomposed for this task in a two stage filtering process in order to derive the required signals for the adaptive approximation and for the residual derivation. As a result, the decoupling of the learning and noise dampening tasks, into two separate and independent tasks, is achieved giving the designer greater flexibility. The second task for fault identification and estimation is accomplished by using learning methods. Specifically, two estimation models are constructed, one for process and one for sensor faults, that utilize adaptive approximation methods to learn the potential fault that has occurred, whilst the fault type is determined on an exclusion-based logic. Finally, fault detectability and identification conditions are derived for both process and sensor faults, that provide an implicit characterization of the faults that can be detected and identified by the proposed scheme and, simulation results indicate the effectiveness of the unified fault diagnosis framework.

1.2 Contribution

The main goal and contribution of this thesis, is the derivation of a fault detection approach relying on filtering so that the effect of measurement noise is dampened allowing the derivation of tight detection thresholds. This is investigated under a wide range of cases consisting all the combinations of continuous/discrete time and the full-state measurements/input-output case. Moreover, in the case of interconnected subsystems, the distributed nature of the approach is further investigated and it is shown to possess important isolation characteristics allowing to derive a high-level fault isolation scheme for process and sensor faults. Finally, the devised filtering framework is exploited so that a unified framework is designed by integrating adaptive approximation methods for learning the modeling uncertainties with filtering for dampening the measurement noise. In this way, the learned modeling uncer-

Christodoulos Kallis

tainty can be used for the residual signal generation, which essentially "cancels out" the true modeling uncertainty, leading to even tighter detection thresholds and enhanced fault detectability. The use of learning is further exploited for identifying the type of the fault that has occurred (process or sensor) and generating an estimation of the fault.

The organization of the dissertation is as follows:

Chapter 2 conducts a literature review on the topic of fault diagnosis.

Chapter 3 proposes a distributed fault filtering approach for the detection of faults, in a class of interconnected, nonlinear, continuous-time systems with modeling uncertainty, measurement noise and disturbances. As a result, a robust fault detection scheme is designed with the state measurements, which are corrupted by measurement noise, being filtered by p-th order filters. Essentially, filtering dampens the effect of measurement noise in a certain frequency range allowing to set less conservative detection thresholds and thus enhancing fault detectability. This chapter also includes a rigorous detectability analysis providing results regarding the magnitude of the detected faults, an upper bound on the detection time and the relation of the detection time with respect to the order and pole locations of the filters used. Simulation results demonstrate the effectiveness of the proposed distributed fault filtering approach.

The chapter is based on the following publications [49], [48]:

- C. Keliris, M. M. Polycarpou, and T. Parisini, "A Distributed Fault Detection Filtering Approach for a Class of Interconnected Continuous-Time Nonlinear Systems," IEEE Transactions on Automatic Control, vol. 58, no. 8, pp. 2032-2047, Aug. 2013.

- C. Keliris and M. M. Polycarpou, "A Distributed Fault Detection Filtering Approach for a Class of Interconnected Continuous-Time Nonlinear Systems," in Proc. of 50th IEEE Conference on Decision and Control and European Control Conference, Orlando, 2011, pp. 89-94.

Chapter 4 develops a nonlinear observer-based approach for distributed fault detection of a class of interconnected input-output nonlinear systems, which is robust to modeling uncertainty and measurement noise. First, a nonlinear observer design is used to generate the residual signals required for fault detection. Then, a distributed fault detection scheme and the corresponding adaptive thresholds are designed based on the observer characteristics and, at the same time, filtering is used in order to attenuate the effect of measurement noise, which facilitates less conservative thresholds and enhanced robustness. Hence, the proposed

Christodoulos Keliris

distributed robust fault detection scheme is inherently tied with the observer performance. Finally, under suitable conditions, a simplified estimation model is formulated and a modified distributed fault detection filtering scheme is proposed. In both cases, the respective fault detectability condition characterizing quantitatively a class of detectable faults is derived and simulation results are used to illustrate the performance of the proposed distributed fault detection approach.

The results presented in this part are based on the following publications [52], [50]:

- C. Keliris, M. M. Polycarpou, and T. Parisini, "A Robust Nonlinear Observer-based Approach for Distributed Fault Detection of Input-Output Interconnected Systems," Automatica, vol. 53, no. 3, pp. 408-415, 2015.

- C. Keliris, M. M. Polycarpou, and T. Parisini, "A Distributed Fault Detection Filtering Approach for a Class of Interconnected Input-Output Nonlinear Systems," in Proc. of European Control Conference, Zurich, 2013, pp. 422-427.

Chapter 5 develops a discrete-time filtering framework for distributed fault detection in a class of interconnected, nonlinear, discrete-time systems with modeling uncertainties and measurement noise under full state measurement. The design of the discrete-time framework allows for direct implementation in real-world applications as required by digital computers and microprocessors. Moreover, in this chapter some practical issues arising such as the IIR/FIR filters required for the generation of the threshold signals are investigated and the fault detectability condition that characterizes quantitatively the class of detectable faults is obtained. Finally, the effectiveness of the approach is illustrated through a simulation example.

The results presented in this part rely on the following publication [5]:

- F. Boem, R. M. G. Ferrari, C. Keliris, T. Parisini, and M. M. Polycarpou, "A Distributed Networked Approach for Fault Detection of Large-scale Systems," IEEE Transactions on Automatic Control (under review).

Chapter 6 develops a distributed fault diagnosis scheme able to deal with process and sensor faults in an integrated way for a class of interconnected input-output nonlinear uncertain discrete-time systems. A robust distributed fault detection scheme is designed, where each interconnected subsystem is monitored by its respective fault detection agent, and according to the decisions of these agents, further information regarding the type of the fault

Christodoulos Keliris

can be deduced. As it is shown, a process fault occurring in one subsystem can only be detected by its corresponding detection agent whereas a sensor fault in a subsystem can be detected by either its corresponding detection agent or the detection agent of another subsystem that is affected by the subsystem where the sensor fault occurred. This discriminating factor is exploited for the derivation of a high-level isolation scheme. Moreover, process and sensor fault detectability conditions characterizing quantitatively the class of detectable faults are derived. Finally, a simulation example is used to illustrate the effectiveness of the proposed distributed fault detection scheme.

The results presented in this part rely on the following publication [53]:

- C. Keliris, M. M. Polycarpou and T. Parisini, "Distributed Fault Diagnosis for Process and Sensor Faults in a Class of Interconnected Input-Output Nonlinear Discrete-Time Systems", International Journal of Control, 2015.

Chapter 7 develops a unified filtering and adaptive approximation based approach for fault diagnosis of process and sensor faults in a class of continuous-time nonlinear systems with modeling uncertainties and measurement noise. The proposed approach integrates learning with filtering techniques and allows the derivation of tight detection thresholds, which is accomplished in two ways: by learning the modeling uncertainty through adaptive approximation methods and, by using filtering for dampening measurement noise. The required signals for both tasks are derived through a two-stage filtering process, by exploiting the properties of the filtering framework. Upon the detection of a fault, two estimation models, one for process and one for sensor faults, are initiated in order to identify the type of the fault. Each estimation model, utilizes learning to estimate the potential fault that has occurred and, adaptive identification thresholds for each estimation model are designed. The fault type is deduced on an exclusion-based logic, by identifying it as the alternate of the one for which the identification threshold is exceeded. Rigorous, fault detectability and identification conditions for process and sensor faults are derived characterizing quantitatively the class of faults that can be detected and identified by the proposed scheme. Finally, simulation results are used to demonstrate the effectiveness of the proposed approach.

The results presented in this part rely on the following publications [54], [51]:

- C. Keliris, M. M. Polycarpou and T. Parisini, "A Unified Fault Diagnosis Approach Utilizing Filtering and Adaptive Approximation for Process and Sensor Faults in a Class of Continuous-Time Nonlinear Systems", (to be submitted in IEEE Transactions on Neural Networks and Learning Systems).

Christodoulos Keliris

- C. Keliris, M. M. Polycarpou and T. Parisini, "A Distributed Fault Diagnosis Approach Utilizing Adaptive Approximation for a Class of Interconnected Continuous-Time Nonlinear Systems," in Proc. of Control and Decision Conference, Los Angeles, 2014, pp. 6536–6541.

Chapter 8 provides some concluding remarks and possible future research directions.

Christodoulos Keliris

Chapter 2

Literature review

The smooth operation and reliability of large-scale interconnected systems is a key requirement in the modern technological world. Examples of such systems include manufacturing systems, chemical plants and critical infrastructure systems such as telecommunication networks, electric power systems and water distribution networks. The need for early detection of faults before they lead to major failures is of crucial importance. The problem of Fault Detection and Isolation (FDI) is not new and there are many important survey papers [32, 37, 42, 84–86] and books [4, 10, 46, 64] on this topic.

Generally, fault diagnosis comprises of several steps [37]: 1) *Fault detection* is the indication that something is wrong. 2) *Fault isolation* is the determination of the location of the fault. 3) *Fault identification* is the determination of the size of the fault. 4) *Fault accommodation* is the needed control reconfiguration to dampen the effect of faults.

A control system comprises mainly of three parts: the actuators, the plant components and the sensors, therefore a fault may appear in any of these. In fact, a fault alarm may be triggered either due to a correct detection of a fault to the aforementioned components or due to a false detection of a fault which is a consequence of modeling inaccuracies, measurement noise and weakened robustness. Apart from the fault origination we can further distinguish the fault modes that can occur to either abrupt faults or incipient faults. *Abrupt faults* are sudden, step-like changes that appear almost instantaneously and can lead to immediate component or even general system failure. On the other hand, *incipient faults* are slowly developing faults that occur due to parameter changes of the components because of their continuous operation and diminishing lifetime. These changes develop slowly and are initially small, thus harder to detect. These sort of faults are better prevented at the maintenance stage where early identification of potentially damaged components may prolong the

Christodoulos Keliris

life of the system and prevent any future failures.

The faults are often classified as process, sensor or actuator faults as indicated in Figure 2.1. Process faults alter the dynamics of the system, sensor faults alter the measurement readings and actuator faults modify the controllers' influence on the system.

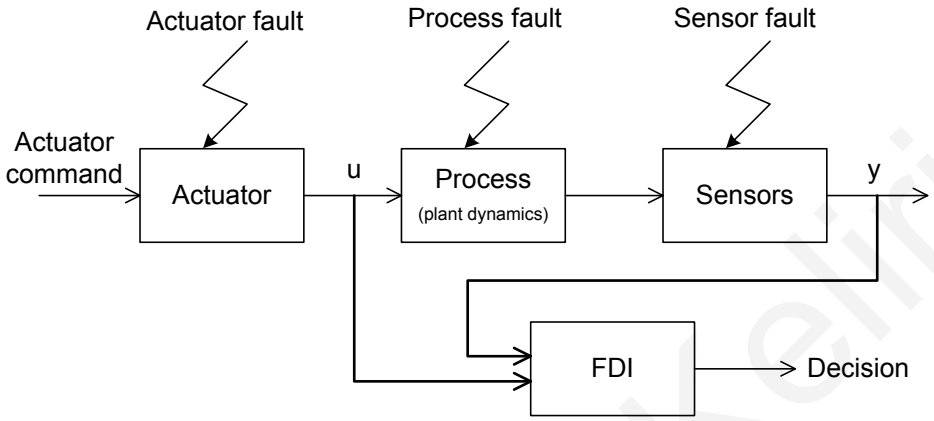


Figure 2.1: Fault types and FDI.

Furthermore, faults can also be classified in the following categories based on the way they act on the system:

a) Additive measurement faults, which are minor differences between the measured signal and its actual value, e.g. sensor biases.

b) Additive process faults, which are unknown inputs (disturbances) that act on the system and affect the system output independently of the measured input, e.g. plant leaks (of course, a leak that is depended on pressure would require nonlinear modeling to be represented more accurately).

c) Multiplicative process faults which are changes in the parameters of the dynamics of the system and are developing either slowly or abruptly. They are usually a result of the diminishing lifetime of the plant components.

d) Additive actuator faults which alters the effect of the controller on the system.

As control systems are becoming more and more complex there is a growing demand for more robust and fault tolerant realizations. Therefore the Fault Detection Isolation and Accommodation (FDIA) concept becomes inherently essential to any modern system. There are mainly two methods in which redundancy can be achieved. At first, there is the *physical redundancy* (or hardware redundancy) approach which requires multiple physical copies of the component or subsystem and it works with the help of a majority voting scheme. This

Christodoulos Keliris

approach though can have high production and maintenance costs, requires greater physical space and, more importantly, it may increase the probability of fault occurrence due to the increased number of physical components. On the other hand, there is the *analytical redundancy* approach which is based on a mathematical model of the system. In this way, the actual physical signals that are measured, are compared to the respective signals given by the mathematical model of the process and their difference constitutes the *residuals*. From thereon, the residuals are utilized by powerful information processing techniques that can indicate the occurrence of faults. Of course, this approach has a high computational burden and is sensitive to false alarms since the modeling uncertainties of the mathematical model may be mistakenly passed as faults.

2.1 General procedure of model-based FDI

The FDI scheme comprises of two steps: initially the generation of residuals and afterwards the use of these residuals in order to identify and isolate any potential faults. At the first step of *residual generation*, the actual system signals are compared to the respective signals that are given by the mathematical model and their difference leads to the residual vector. Under the ideal conditions of no faults, no modeling uncertainties and no measurement noise and disturbances the residual vector should be zero, something that does not occur in practice. The second step which is also called the *decision maker* is the most crucial one as it is the one that will form a decision whether a fault has occurred or not. If the decision is positive then further analysis is conducted to identify the faults' type, location and even perhaps its size. The evaluation of residuals at this decision stage usually involves threshold testing or statistical tests. Sometimes, fault association can be defined based on various fault signatures that are available beforehand enabling easier fault detection and isolation. It should also be noted that whilst a single residual signal is sufficient to detect faults, a set of residuals is usually required for fault isolation [65].

2.2 Residual generation methods based on analytical redundancy

There are various methods to generate the residual vector under the analytical redundancy framework but they can generally be divided in two approaches: the state estimation

Christodoulos Kellis

techniques (such as parity space approach, observer schemes and detection filters) and the parameter identification techniques. Moreover, to ease the fault isolation the residuals can be enhanced during their generation so that they contain specific isolation properties. The main residual enhancement techniques are the generation of structured and directional residuals [38, 85]. In the *structured residuals* scheme, each fault affects a specific subset of the residuals and any residual responds only to a specific subset of faults. Therefore, due to the dependence of the residuals on the faults, certain patterns appear on the residual vector that can be used for fault isolation. In the *directional residuals* scheme, each fault amounts to a specific direction in the residual space thus fault isolation is concluded by selecting the direction that the generated residual vector lies the closest to. More information regarding these techniques can be found in the books by Gertler [39] and Isermann [46]. In the literature many methods have been proposed for the generation of residuals most of which their concept falls in one of the following approaches:

- *Parity space approach*. Essentially the method checks the consistency of the mathematical equations by using the actual measurements and a fault is declared whenever predetermined error thresholds are exceeded. Further information can be found in [38] and the references therein.

- *Observer schemes*. *Fault detection filter* (FDF) is the first kind of observer-based residual generators proposed by Beard and Jones in the early 70's. Another, and in fact one of the most frequently used, method is the *Diagnostic Observer* approach. According to this approach and in the case of deterministic systems, observers are being used to reconstruct the systems' outputs $\hat{y}$ from measurements and, the residual is simply the output estimation error $e = y - \hat{y}$. In the case of stochastic systems where noise is present, the observers are substituted with Kalman filters and the residual is the innovation which under the fault free case should be white noise with zero mean and known covariance. The isolation of faults can be enhanced with the use of a bank of residual generators under the *Dedicated Observer Scheme (DOS)* or the *Generalized Observer Scheme (GOS)* [32, 33]. The Dedicated Observer Scheme was initially developed by Clark [17] to isolate sensor faults. It assumes that N faults have to be detected and isolated, thus it constructs N residual generators (observers) each of which is driven by only one output. Therefore each residual is affected by only one fault which can be isolated. The Generalized Observer Scheme was developed by Frank [31] and it also assumes that N faults have to be detected and isolated. Again, it constructs N

Christodoulos Kallis

observers from which the i-th observer is driven with all but one measured variable (except y_i). Therefore, it can isolate a single fault at a time, whereas the DOS scheme can isolate multiple concurrent faults. A major drawback of the DOS scheme is that its application is limited to sensor faults something that is not the case with the GOS scheme which can always be applied.

- *Parameter estimation*. This method is particularly suited to the detection of incipient faults and it is extensively studied in the survey papers by Isermann [44] and Frank [32] and the books by Patton *et al.* [65] and Isermann [46]. Using system identification methods (utilizing the input and output signals), the parameters of a mathematical model of the system can be obtained (recursively and on-line) across different time intervals and compared to their respective values based on a nominal model. Any significant difference could indicate the occurrence of a fault and, a relation between parameter changes and faults can be formed with the use of pattern recognition methods.

The main issue with the analytical redundancy approach is that a mathematical model of the system is never known exactly, and thus, the resulting residual vector is never identically zero. Error bounds can be defined, but still the proper choice remains a major problem. If bounds are chosen too narrow it may lead to false alarms. Too wide and faults may pass initially unnoticed. Discussion of the proper threshold choice can be found in the paper by Emami-Naeini *et al.* [23]. As a result, there is a growing demand for robust residual generation to reduce the sensitivity of the residual against the effect of modeling errors, noise and disturbances. This issue can be tackled either by the use of enhanced techniques for robust residual generation or by choosing appropriately the level of the error threshold which can also change adaptively as discussed in the book by Patton *et al.* [65]. The need for an accurate mathematical model can be overcome by making use of qualitative approaches to aid fault diagnosis, such as classification techniques and inference methods. Further information regarding these approaches can be sought in [44–46, 58].

2.3 FDI in nonlinear systems

So far, the methods that have been mentioned apply mainly to linear systems and their application to nonlinear systems requires model linearization around the operating point. Of course, some of the methods were extended to the nonlinear systems case. Such example

Christodoulos Keliris

is the design of observers, for which an excellent treatment for various cases can be found in [34, 35, 41]. In fact, the observer approach is one of the main methods to address the fault diagnosis problem [33, 61, 64]. The basic idea of this approach relies in reconstructing the outputs of the system from the measurements using observers in the case of deterministic systems or Kalman filters in the case of stochastic systems. The estimation error (the difference between the measurement and the predicted measurement from the observer or Kalman filter) is then used as a residual for the detection and isolation of faults. The residual signal that is obtained includes not only potential fault signals, but also the combined effect of the modeling uncertainty, the process and the measurement noise. One way to overcome the problems created by noise is by using Kalman filter residual generators in the case of linear stochastic systems or Extended Kalman Filters in the case of nonlinear stochastic systems. The use of Kalman filters leads to some statistical testing of the residuals, usually by assuming that the noise is a zero-mean process while the faults are more deterministic in nature [2]. Some Kalman state estimators have been specially designed for FDI of stochastic systems such as multiple model filters [18, 60, 62]. Other observer designs targeting specific forms of nonlinearities have been proposed; i.e. works for bilinear nonlinearity can be found in Dingli *et al.* [91] and observers based on differential geometric methods for fault-affine model forms can be found in Yang *et al.* [89]. Other approaches for bilinear systems can be found in [90,92], for nonlinear systems with geometric methods in [20] and a combination of bilinear systems with geometric methods in [40]. Essentially, bilinear systems are a special case of more general nonlinear systems and many physical systems can be represented in this way [34].

As pointed out in [33], the observers used in fault diagnosis are primarily output observers which simply reconstruct the measurable part of the state variables, rather than state observers which are required for control purposes. The use of state observers for nonlinear systems has not been used extensively for the FDI problem, even though analytical results regarding the stability of the nonlinear observers and design procedures have been established [3, 36, 47, 56, 72, 76, 81, 102]. The main issue with the observer approach is that the design of observers for nonlinear systems with asymptotically stable error dynamics is not an easy task even when the nonlinearities are fully known. As a result, the research in fault diagnosis for nonlinear systems utilizing state observers is more limited [1, 35, 41, 73, 93].

In general, the research that has been conducted for nonlinear systems is more limited [10, 30, 35, 41, 70, 94] in comparison to the linear systems case. In other works, Chen & Saif [14] and Pertew *et al.* [67] extend the unknown input observer approach to a class

Christodoulos Kallis

of Lipschitz nonlinear systems, De Persis & Isidori [19] tackle the FDI problem using a differential-geometric approach for a class of nonlinear systems while Zhang *et al.* [96] develop an FDI scheme for nonlinear systems utilizing a learning methodology. A different approach for fault diagnosis that has been developed is under a stochastic framework and relevant work has been conducted by Basseville [2], Castillo *et al.* [9] and Zufiria [103].

An important issue that is often overseen in the literature, is the presence of measurement noise. In most real world applications such uncertainty may influence significantly the performance of fault detection schemes by causing, for example, false alarms or missed faults (if the detection threshold is selected too high). In the case of nonlinear systems with noisy measurements little work has been conducted. In Zhang *et al.* [95] sensor bias fault isolation in a class of nonlinear systems is conducted where some measurement uncertainty is present. More recently, Ferrari *et al.* [29] investigated the problem of distributed fault detection and isolation for large-scale, discrete-time, nonlinear, uncertain systems with measurement uncertainty where the monolithic system is decomposed into several smaller interconnecting subsystems and in Boem *et al.* [6] the proposed methodology was presented in the case of continuous-time systems. In Reppa *et al.* [74] an adaptive approximation is developed for multiple sensor fault detection and isolation, in Ferdowsi *et al.* [25] an online outlier identification and removal scheme for improving the fault detection performance is proposed and, in Talebi *et al.* [80] a recurrent neural-network-based sensor and actuator fault detection and isolation for nonlinear systems with application to the satellite's attitude control subsystem is considered. The treatment for the measurement uncertainty in the aforementioned works though, relies on having bounding functions for the measurement uncertainty in order to guarantee no false alarms whilst no action is taken to alleviate its effects. The problem of FDI in the nonlinear systems case that tackle the measurement noise, has been dealt by De Persis & Isidori [20] by utilizing a fault detection filter with game-theoretic-optimal sensitivity which attenuates the effect of the measurement noise on the residual signal while at the same time it retains the effect of the fault on the residual. Similar work in the case of linear systems has been conducted by Chung & Speyer [16] and Chen & Speyer [11], [12]. Proposed schemes for dealing with the measurement noise in the case of linear systems has also been presented by Zhang & Jaimoukha [100, 101]. An excellent recent survey paper that discusses various model-based techniques to generate residuals that are robust to noise, unknown disturbance, and model uncertainties, as well as various statistical techniques of testing the residuals for faults can be found in Hwang *et al.* [42]. The main reason for the reduced research in dealing with the measurement noise under a deterministic framework,

Christodoulos Kellis

is that usually process disturbances and measurement noise are modeled under a stochastic framework, and their effect to the residual signals is considered by using statistical decision methods [2] and, this work aims primarily in filling this gap.

2.4 Departure from centralized approaches

So far, the previous discussion concerns mostly centralized approaches for the fault diagnosis problem, where information about the state of the system are gathered and processed centrally. Recent advances in communications and distributed sensing motivated the departure from centralized fault diagnosis approaches [4, 10, 32, 64, 85] towards the development of hierarchical, decentralized and distributed schemes [6, 26, 28, 29, 55, 57, 66, 79, 87, 98]. From a practical perspective, gathering the distributed information into a central processing station to apply a centralized approach for the fault diagnosis task is counter-productive due to communication overload and the requirement for higher computational power. Moreover the processing of the information at a centralized station imposes several risks since the station constitutes a *single point of failure*, thus deteriorating the fault diagnosis capability. On the other hand, distributed systems either by nature or through the decomposition of large scale systems into smaller, interconnected subsystems provide a solid base for developing fault diagnosis approaches with inherent robustness.

In many cases, a distributed FDI framework is not an option but a necessity, since many factors contribute to this formulation such as the large scale nature of the system to be monitored, its spatial distribution, the inability to access certain parts of the system from a remote monitoring component and therefore local diagnosis should be performed. Generally, in the research literature, there is no clear indication of quantitative benefits of the various schemes (i.e. centralized, decentralized, distributed) over the other. Qualitatively though, the main benefits of using a distributed fault diagnosis scheme can be summarized as follows: a) enhanced safety robustness, since centralized approaches are subject to single-point-of-failure, b) less communication bandwidth requirements since only necessary information are exchanged rather than conveying all the information in a single location as in centralized approaches, c) less computation power requirements, d) scalability benefits; the distributed scheme allows for more flexibility in adding subsystems with respective fault detection modules without requiring any computational modifications in the already existing architecture.

Christodoulos Kallis

2.5 Process and sensor faults

In the research literature, there is the tendency of dealing with the problem of fault diagnosis for process and sensor faults separately and not simultaneously, something that poses a significant limitation for real world application. For example, in the problem of fault diagnosis for process faults the sensors are considered healthy. But, apart from erroneous detection results a faulty sensor may lead, it can also lead to degraded tracking or regulation performance or even endanger the stability of control systems. Acknowledging, that sensors are prone to faults and utilizing sensor validation approaches is crucial in the overall system stability and reliability. Similarly, in the problem of fault diagnosis for sensor faults it is assumed that there are no process faults. Obviously, dealing with the process and sensor fault problem separately incurs the danger of false alarms due to monitoring the specific fault type and ignoring the other, something that results in unnecessary component replacement and increased maintenance costs.

The research conducted on the fault diagnosis problem that deals with process and sensor faults simultaneously is very limited [22, 75, 93, 97, 99]. In [75] the sensor and process fault detection problem is addressed using multi sensor data fusion techniques based on based on the adaptive extended Kalman filter algorithm, whereas in [22] a unified framework for dealing with joint diagnosis of process and sensor faults is proposed along with fault identification and reconstruction via principal component analysis. In the context of analytical redundancy methods, [97] develops a fault isolation approach to determine which process or sensor fault, among two respective fault classes, has occurred. Based on the assumption that only a single fault occurs (either a process or a single sensor), adaptive approximation methods are used in order to build a fault detection estimator and suitable fault isolation estimators that correspond to the process and sensor faults that are able to determine which fault has occurred. Moreover, the work in [99] follows along similar concepts of [97]. In [80] a recurrent neural-network based fault detection scheme for nonlinear systems is proposed, which employs two nonlinear-in-parameters neural networks to isolate actuator and sensor faults; the fault determined when the output of one of the neural networks produces a non-zero output indicating the faulty condition. In [82] a nonlinear observer-based fault diagnostics scheme, dealing with process and sensor faults, for nonlinear systems in discrete time is proposed. The scheme consists of an artificial immune system as an online approximator, which identifies the fault type by monitoring the outputs' magnitude of the two online approximators (state and output) as in [80]. Finally, in [93] a distributed detection scheme for process

Christodoulos Kallis

and sensor faults for a class of input-output interconnected systems under continuous time is proposed, but the estimator design is conducted under some potentially restrictive conditions and, deals only with the fault detectability issue.

Christodoulos Keliris

Chapter 3

A Fault Detection Filtering Approach

The primary objective of this chapter is the devise of a suitable filtering framework for the detection of faults and the conduction of a thorough detectability analysis, in a class of interconnected, nonlinear, continuous-time systems with modeling uncertainty, measurement noise and disturbances. As a result, a robust fault detection scheme is designed with the state measurements, which are corrupted by measurement noise, being filtered by p-th order filters. Essentially, filtering dampens the effect of both measurement noise and disturbances in a certain frequency range allowing to set less conservative detection thresholds and thus enhancing fault detectability. At the same time though, it imposes some new challenges such as how to properly design the estimator and derive the corresponding thresholds. A key novelty of this work is that the residual and threshold signals are constructed by embedding the filters into the design in a way that takes advantage of the filtering noise suppression properties.

By assuming that the filtering dampens the measurement noise and disturbances, adaptive fault detection thresholds are obtained under a rigorous analytical framework guaranteeing no false alarms. Further on, a fault detectability condition is devised that characterizes the class of detectable faults. The distributed fault detection scheme is based on local fault detection agents with each one assigned to monitor one subsystem. Each local fault detection agent receives the input and output measurements of the subsystem it monitors and also the output measurements of all the interconnected subsystems that influence the subsystem under consideration. Finally, the local fault detection scheme provides a decision regarding the health of the subsystem it monitors. The implementation of the scheme is presented in detail, along with practical issues and potential solutions. Further on, by considering a specific class of filters taking on the form $H_p(s) = \frac{\alpha^p}{(s+\alpha)^p}$ and by simplifying the assumption related to the

Christodoulos Kellinis

modeling uncertainty, additional analytical results are derived providing more information regarding the properties of the proposed filtering scheme. These additional results characterize the class of detectable faults with respect to their magnitude and establish an upper bound on the detection time, as well as the relation between the detection time with the order p of the filter and its pole location α, thus shedding intuition into the mathematical complexity.

The chapter is organized as follows: in Section 3.1, the problem of distributed fault detection of a class of nonlinear dynamical systems with modeling and measurement uncertainties and process noise dealt with is formulated. In Section 3.2, the design of the distributed fault detection scheme based on a filtering approach is presented in detail and the implementation of the proposed scheme is demonstrated along with proposed solutions to some practical issues. In Section 3.3, the fault detectability condition characterizing the class of faults detectable by the proposed methodology is derived and the fault detectability is further investigated by considering a special case for the filter. In Section 3.4, results regarding the detection time in the special case mentioned before are derived and in Section 3.5 a simulation example illustrates the concepts presented through different case scenarios showing the impact of noise on the fault detection problem and emphasizing the effectiveness of the proposed scheme. Finally, Section 3.6 provides some concluding remarks.

3.1 Problem Formulation

Consider a large-scale distributed nonlinear dynamic system made of N subsystems Σ_I, $I \in \{1, ..., N\}$, each of which is described by the differential equation:

$$\Sigma_I : \begin{cases} \dot{x}_I = f_I(x_I, u_I) + g_I(x_I, \bar{x}_I, u_I) + \zeta_I(t) \\ \qquad +\eta_I(x_I, \bar{x}_I, u_I, t) + \beta_I(t - T_0)\phi_I(x, u_I) & (3.1) \\ y_I(t) = x_I(t) + \xi_I(t), & (3.2) \end{cases}$$

where $x_I \in \mathbb{R}^{n_I}$, $u_I \in \mathbb{R}^{m_I}$ and $y_I \in \mathbb{R}^{n_I}$ are the state, input and measured output vectors of the I-th subsystem respectively, $x \triangleq \left[x_1^T, x_2^T, \ldots, x_N^T\right]^T \in \mathbb{R}^n$ is the state vector of the overall system, $\bar{x}_I \in \mathbb{R}^{\bar{n}_I}$ contains the state variables of the other subsystems $J \in \{1, \ldots, N\} \setminus \{I\}$ that affect the I-th subsystem, $f_I : \mathbb{R}^{n_I} \times \mathbb{R}^{m_I} \mapsto \mathbb{R}^{n_I}$ is the known local (nominal) function dynamics of the I-th subsystem and $g_I : \mathbb{R}^{n_I} \times \mathbb{R}^{\bar{n}_I} \times \mathbb{R}^{m_I} \mapsto \mathbb{R}^{n_I}$ is the known part of the interconnection function between the I-th and the other subsystems. The vector function $\eta_I : \mathbb{R}^{n_I} \times \mathbb{R}^{\bar{n}_I} \times \mathbb{R}^{m_I} \times \mathbb{R}^+ \mapsto \mathbb{R}^{n_I}$ is the overall modeling uncertainty associated with the nominal and the interconnection functions, $\zeta_I(t) \in \mathbb{R}^{n_I}$ is the unknown disturbance associated with the I-th subsystem and $\xi_I \in \mathcal{D}_{\xi_I} \subset \mathbb{R}^{n_I}$ ($\mathcal{D}_{\xi_I}$ is a compact

Christodoulos Keliris

set) represents the measurement noise. The state vectors x_I, $I \in \{1, ..., N\}$ are considered unknown whereas their noisy counterparts y_I are known. The term $\beta_I(t - T_0)\phi_I(x, u_I)$ characterizes the fault function dynamics affecting the I-th subsystem including its time evolution. More specifically, the term $\phi_I : \mathbb{R}^n \times \mathbb{R}^{m_I} \mapsto \mathbb{R}^{n_I}$ is the unknown fault function and the term $\beta_I(t - T_0) : \mathbb{R} \mapsto \mathbb{R}^+$ denotes the time evolution of the fault, where T_0 is the unknown time of the fault occurrence [70]. Note that the fault function ϕ_I may depend on the global state variable vector x and not only on the local state vector x_I. From a practical perspective, this allows for a more general class of faults compared to the case where the fault ϕ_I is a function of x_I only. In addition, this allows for propagative fault effects to be transferred across neighboring subsystems (as it is the case in real networks such as electric power systems, transportation systems, etc). The fault time profile $\beta_I(t - T_0)$ can be used to model abrupt faults or incipient faults using a decaying exponential type function:

$$\beta_I(t - T_0) \triangleq \begin{cases} 0 & \text{if } t < T_0 \\ 1 - e^{-b_I(t-T_0)} & \text{if } t \geq T_0 \end{cases} \tag{3.3}$$

where $b_I > 0$ is typically an unknown parameter which denotes the fault evolution rate. Abrupt faults correspond to the limit $b_I \to \infty$: in this case, the time profile $\beta_I(t - T_0)$ becomes a step function. In general, small values of b_I indicate slowly developing faults (incipient faults) whereas large values of b_I make the time profile $\beta_I(t - T_0)$ approach a step function (abrupt faults).

The objective is to design and analyze a distributed fault detection scheme, with each subsystem Σ_I being monitored by a local fault detection module that receives local measurements and partial information from interconnected subsystems.

In general, the distributed fault detection scheme is composed of N local filtered fault detection modules $\mathcal{F}_I$, one for each subsystem Σ_I. Each local fault detection module $\mathcal{F}_I$ requires the input and output measurements of the subsystem Σ_I that is monitoring and also the measurements of all interconnecting subsystems Σ_J that are influencing Σ_I. Note that these last measurements are communicated by neighboring fault detection modules $\mathcal{F}_J$, and not by the subsystems Σ_J. Therefore, there is the need of communication between the fault detection modules depending on their interconnections. Note that, the information exchanged among the subsystems is readily available since it is constituted by quantities $\bar{x}_I$ that are measurable with some uncertainty as $\bar{y}_I = \bar{x}_I + \bar{\xi}_I$ (the noisy counterpart of $\bar{x}_I$). Therefore, the distributed nature of the scheme stems from the fact that there is communication between the fault detection modules depending on their interconnections. The example shown in Fig-

Christoforos Kellis

ure 3.1 illustrates the distributed fault detection scheme for the case of three subsystems Σ_1, Σ_2, Σ_3, where Σ_1 affects Σ_2, and Σ_2 affects Σ_3.

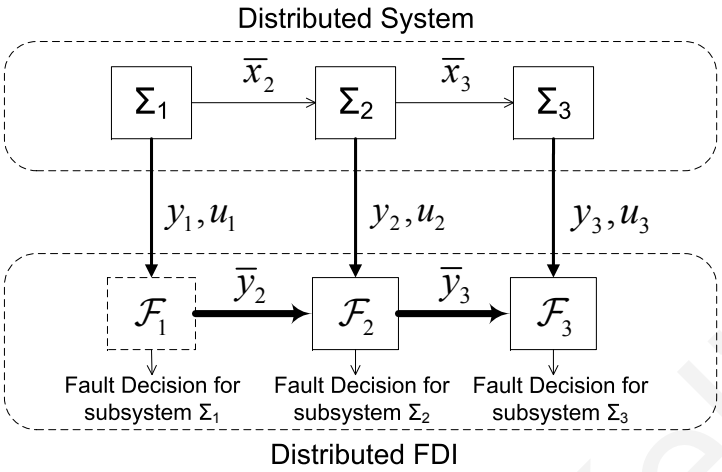


Figure 3.1: Distributed fault detection scheme for the case of three subsystems where Σ_1 affects Σ_2, and Σ_2 affects Σ_3.

It is assumed that there exist feedback controllers yielding a local control action u_I such that some desired control objectives are achieved. In this work, we do not deal explicitly with the control problem, but instead, we consider the design and analysis of fault detection algorithms in the presence of faults ϕ_I, modeling uncertainties η_I, disturbances ζ_I and measurement noise ξ_I. Particular emphasis is placed on addressing issues related to measurement noise and disturbances, which may affect significantly the performance of the fault detection scheme.

The notation $|\cdot|$ used in this chapter indicates the absolute value of a scalar function. In addition, the notation $y(t) = H(s)\big[x(t)\big]$ denotes the output $y(t)$ of any signal $x(t)$ which is passed through a filter with transfer function $H(s)$. In other words, $y(t)$ is the output of a linear system represented by the transfer function $H(s)$ with $x(t)$ as input. In terms of more rigorous notation, let $h(t)$ be the impulse response associated with $H(s)$; i.e. $h(t) \triangleq \mathcal{L}^{-1}[H(s)]$, where $\mathcal{L}^{-1}$ is the inverse Laplace transform operator. Then $y(t) = H(s)\big[x(t)\big] = \int_0^t h(\tau)x(t-\tau)\,\mathrm{d}\tau$. It must be stressed that the simplified notation $H(s)\big[x(t)\big]$ is used extensively in the adaptive control literature.

The following assumptions are used throughout the chapter:

***Assumption* 3.1.** *For each subsystem Σ_I, $I \in \{1, ..., N\}$, the local state variables $x_I(t)$ and the local inputs $u_I(t)$ belong to a known compact region $\mathcal{D}_{x_I}$ and $\mathcal{D}_{u_I}$ respectively before and after the occurrence of a fault, i.e. $x_I(t) \in \mathcal{D}_{x_I}$, $u_I(t) \in \mathcal{D}_{u_I}$ for all $t \geq 0$.*

Christodoulos Keliris

***Assumption* 3.2.** *The modeling uncertainty $\eta_I^{(k)}$ (k denotes the k-th component of η_I) in each subsystem is an unstructured and possibly unknown nonlinear function of x_I, $\bar{x}_I$, u_I and t but uniformly bounded by a known positive function $\bar{\eta}_I^{(k)}$, i.e.,*

$$|\eta_I^{(k)}(x_I, \bar{x}_I, u_I, t)| \leq \bar{\eta}_I^{(k)}(y_I, \bar{y}_I, u_I), \quad k = 1, 2, \ldots, n_I \tag{3.4}$$

for all $t \geq 0$ and for all $(x_I, \bar{x}_I, u_I) \in \mathcal{D}_I$, where $\bar{y}_I = \bar{x}_I + \bar{\xi}_I$ is the measurable noisy counterpart of $\bar{x}_I$, $\bar{\xi}_I \in \mathcal{D}_{\bar{\xi}_I} \subset \mathbb{R}^{\bar{n}_I}$ and $\bar{\eta}_I^{(k)}(y_I, \bar{y}_I, u_I) \geq 0$ is a known bounding function in some region of interest $\mathcal{D}_I = \mathcal{D}_{x_I} \times \mathcal{D}_{\bar{x}_I} \times \mathcal{D}_{u_I} \subset \mathbb{R}^{n_I} \times \mathbb{R}^{\bar{n}_I} \times \mathbb{R}^{m_I}$. The regions $\mathcal{D}_{\bar{\xi}_I}$ and $\mathcal{D}_I$ are compact sets.

Assumption 3.1 is required for well-posedness since in this work we do not address the control design and fault accommodation problem, but instead the fault detection problem. Assumption 3.2 characterizes the class of modeling uncertainties being considered. In practice, the system can be modeled more accurately in certain regions of the state space. Therefore, the fact the bound $\bar{\eta}_I$ is a function of y_I, $\bar{y}_I$ and u_I provides more flexibility by allowing the designer to take into consideration any prior knowledge of the system. Moreover, the bound $\bar{\eta}_I$ is required in order to distinguish the effects between modeling uncertainty and faults. Assumption 3.2 will be simplified later on by assuming further that the bound is constant $\bar{\eta}_I$ in order to analyze some key properties of the proposed filtering approach.

To dampen the effect of measurement uncertainty $\xi_I(t)$, each measured variable $y_I^{(k)}$ (k-th component of y_I) is filtered by $H(s)$, where $H(s)$ is a p-th order filter with strictly proper transfer function

$$H(s) = \frac{\left(d_{p-2}s^{p-2} + d_{p-3}s^{p-3} + \ldots + d_0\right)s}{s^p + c_{p-1}s^{p-1} + \ldots + c_1 s + c_0}. \tag{3.5}$$

The choice of a particular type of filter to be used is application dependent, and it is made according to the available a-priori knowledge on the noise properties. Usually, measurement noise is constituted by high frequency components and therefore the use of low-pass filter for dampening noise is well justified. On other occasions, one may want to focus the fault detectability on a prescribed frequency band of the measurement signals and hence, choose the filter accordingly. A rigorous investigation of the filtering impact according to the poles' location and filters' order on the detection time is presented in Section 3.3.

Generally, each measured variable $y_I^{(k)}(t)$ can be filtered by a different filter. In this chapter, without loss of generality, we consider $H(s)$ to be the same for all the output variables in order to simplify the notation and presentation.

Christodoulos Kallis

Note that the filter $H(s)$ can be written as $H(s) = sH_p(s)$ where

$$H_p(s) = \frac{d_{p-2}s^{p-2} + d_{p-3}s^{p-3} + \ldots + d_0}{s^p + c_{p-1}s^{p-1} + \ldots + c_1 s + c_0}. \tag{3.6}$$

The filters $H(s)$ and $H_p(s)$ are asymptotically stable and hence BIBO stable. Therefore, for bounded measurement noise $\xi_I(t)$ and bounded disturbance $\zeta_I(t)$, the filtered measurement noise $\epsilon_{\xi_I}(t) \triangleq H(s)\left[\xi_I(t)\right]$ and filtered disturbance $\epsilon_{\zeta_I}(t) \triangleq H_p(s)\left[\zeta_I(t)\right]$ are uniformly bounded as follows:

$$|\epsilon_{\xi_I}^{(k)}(t)| \leq \bar{\epsilon}_{\xi_I}^{(k)} \quad k = 1, 2, \ldots, n_I, \tag{3.7}$$

$$|\epsilon_{\zeta_I}^{(k)}(t)| \leq \bar{\epsilon}_{\zeta_I}^{(k)} \quad k = 1, 2, \ldots, n_I, \tag{3.8}$$

where $\bar{\epsilon}_{\xi_I}^{(k)}$, $\bar{\epsilon}_{\zeta_I}^{(k)}$ are known bounding constants. Depending on the noise characteristics, $H(s)$ can be selected to reduce the bound $\bar{\epsilon}_{\xi_I}^{(k)}$.

3.2 Distributed Fault Detection

3.2.1 Residual and Threshold Signals Generation

By locally filtering the output signal $y_I(t)$ we obtain the filtered output $z_I(t)$:

$$\begin{aligned} z_I(t) &= H(s)\left[y_I(t)\right] \\ &= H(s)\left[x_I(t) + \xi_I(t)\right]. \end{aligned} \tag{3.9}$$

By using $\epsilon_{\xi_I}(t) = H(s)\left[\xi_I(t)\right]$, $\epsilon_{\zeta_I}(t) = H_p(s)\left[\zeta_I(t)\right]$ and the fact that $s[x_I(t)] = \dot{x}_I(t) + x_I(0)\delta(t)$ (where $\delta(t)$ is the delta function) we obtain:

$$\begin{aligned} z_I(t) &= H(s)\left[x_I(t)\right] + \epsilon_{\xi_I}(t) \\ &= H_p(s)\left[\dot{x}_I(t)\right] + H_p(s)\left[x_I(0)\delta(t)\right] + \epsilon_{\xi_I}(t) \\ &= H_p(s)\big[f_I\big(x_I(t), u_I(t)\big) + g_I(x_I(t), \bar{x}_I(t), u_I(t)\big) \\ &\qquad + \eta_I\big(x_I(t), \bar{x}_I(t), u_I(t), t\big) + \beta_I(t - T_0)\phi_I\big(x(t), u_I(t)\big)\big] \\ &\quad + \epsilon_{\xi_I}(t) + \epsilon_{\zeta_I}(t) + h_p(t)x_I(0), \end{aligned} \tag{3.10}$$

where $h_p(t)$ is the impulse response of the filter $H_p(s)$, i.e. $h_p(t) \triangleq \mathcal{L}^{-1}\left[H_p(s)\right]$. Based on (3.1), an estimation model $\hat{x}_I(t)$ for $x_I(t)$ under fault-free operation is generated by considering the known components and by ignoring the unknown modeling uncertainty η_I, the faults ϕ_I and by replacing x_I and $\bar{x}_I$ with y_I and $\bar{y}_I$ respectively. Therefore, we have:

$$\hat{x}_I(t) = \int_0^t \left[f_I(y_I(\tau), u_I(\tau)) + g_I(y_I(\tau), \bar{y}_I(\tau), u_I(\tau))\right] \mathrm{d}\tau + y_I(0),$$

Christodoulos Keliris

which can be written in differential form as

$$\dot{\hat{x}}_I = f_I(y_I, u_I) + g_I(y_I, \bar{y}_I, u_I), \tag{3.11}$$

with the initial condition $\hat{x}_I(0) = y_I(0)$.
The corresponding estimation model for $z_I(t)$, denoted by $\hat{z}_I(t)$, is given by

$$\hat{z}_I(t) = H(s)\big[\hat{x}_I(t)\big], \tag{3.12}$$

and using (3.11), $\hat{z}_I(t)$ becomes

$$\hat{z}_I(t) = H_p(s)\big[f_I\big(y_I(t), u_I(t)\big) + g_I\big(y_I(t), \bar{y}_I(t), u_I(t)\big)\big] + h_p(t)y_I(0). \tag{3.13}$$

The local residual error $r_I(t)$ to be used for fault detection is defined as:

$$r_I(t) \triangleq z_I(t) - \hat{z}_I(t). \tag{3.14}$$

This residual constitutes the basis of the fault detection scheme and it is readily computable from equations (3.9), (3.11) and (3.12). A fault in the overall system is said to be detectable (i.e. a detection decision can be made) when $|r_I^{(k)}(t)| > \bar{r}_I^{(k)}(t)$, for at least one component k in any local subsystem Σ_I, where $\bar{r}_I^{(k)}(t)$ is the detection threshold (to be specified later on).

Prior to the fault ($t < T_0$), the local residual error can be written using equations (3.10), (3.13) and (3.14) as:

$$r_I(t) = H_p(s)\left[\chi_I(t)\right] + \epsilon_{\xi_I}(t) + \epsilon_{\zeta_I}(t) \tag{3.15}$$

where the total uncertainty term $\chi_I(t)$ is defined as:

$$\chi_I(t) \triangleq \Delta f_I(t) + \Delta g_I(t) + \eta_I\big(x_I(t), \bar{x}_I(t), u_I(t), t\big), \tag{3.16}$$

$$\Delta f_I(t) \triangleq f_I\big(x_I(t), u_I(t)\big) - f_I\big(x_I(t) + \xi_I(t), u_I(t)\big), \tag{3.17}$$

$$\Delta g_I(t) \triangleq g_I\big(x_I(t), \bar{x}_I(t), u_I(t)\big) - g_I\big(x_I(t) + \xi_I(t), \bar{x}_I(t) + \bar{\xi}_I(t), u_I(t)\big). \tag{3.18}$$

For simplicity, in the derivation of (3.15) the initial conditions $x_I(0) = y_I(0)$ are assumed to be known. If there is uncertainty in the initial conditions (i.e. $x_I(0) \neq y_I(0)$) then that introduces the extra term $h_p(t)(x_I(0) - y_I(0))$ in (3.15) which however converges to zero exponentially (since $h_p(t)$ is exponentially decaying [21]) and thus does not affect significantly the subsequent analysis.

Christodoulos Keliris

By taking bounds on the local residual error, we obtain:

$$\begin{aligned}
|r_I^{(k)}(t)| &= |H_p(s)\left[\chi_I^{(k)}(t)\right] + \epsilon_{\xi_I}^{(k)}(t) + \epsilon_{\zeta_I}^{(k)}(t)| \\
&\leq |H_p(s)\left[\chi_I^{(k)}(t)\right]| + |\epsilon_{\xi_I}^{(k)}(t)| + |\epsilon_{\zeta_I}^{(k)}(t)| \\
&= |\int_0^t h_p(t-\tau)\chi_I^{(k)}(\tau)\,\mathrm{d}\tau| + |\epsilon_{\xi_I}^{(k)}(t)| + |\epsilon_{\zeta_I}^{(k)}(t)| \\
&\leq \int_0^t |h_p(t-\tau)||\chi_I^{(k)}(\tau)|\,\mathrm{d}\tau + |\epsilon_{\xi_I}^{(k)}(t)| + |\epsilon_{\zeta_I}^{(k)}(t)| \\
&\leq \int_0^t |h_p(t-\tau)|\bar{\chi}_I^{(k)}(\tau)\,\mathrm{d}\tau + \bar{\epsilon}_{\xi_I}^{(k)} + \bar{\epsilon}_{\zeta_I}^{(k)}, \qquad (3.19)
\end{aligned}$$

where $\bar{\chi}_I^{(k)}(t)$ is the bound on the total uncertainty term $\chi_I^{(k)}(t)$, i.e.,

$$0 \leq |\chi_I^{(k)}(t)| \leq \bar{\chi}_I^{(k)}(t). \qquad (3.20)$$

Using Assumption 2, the bound $\bar{\chi}_I^{(k)}(t),\ k = 1, 2, \ldots, n_I$ is defined as:

$$\bar{\chi}_I^{(k)}(t) \triangleq \overline{\Delta f}_I^{(k)} + \overline{\Delta g}_I^{(k)} + \bar{\eta}_I^{(k)}\big(y_I(t), \bar{y}_I(t), u_I(t)\big), \qquad (3.21)$$

where

$$\overline{\Delta f}_I^{(k)} \triangleq \sup_{\substack{(x_I,u_I)\in\mathcal{D}_{x_I}\times\mathcal{D}_{u_I} \\ \xi_I\in\mathcal{D}_{\xi_I}}} |f_I^{(k)}\big(x_I, u_I\big) - f_I^{(k)}\big(x_I+\xi_I, u_I\big)| \qquad (3.22)$$

$$\overline{\Delta g}_I^{(k)} \triangleq \sup_{\substack{(x_I,\bar{x}_I,u_I)\in\mathcal{D}_I \\ (\xi_I,\bar{\xi}_I)\in\mathcal{D}_{\xi_I}\times\mathcal{D}_{\bar{\xi}_I}}} |g_I^{(k)}\big(x_I, \bar{x}_I, u_I\big) - g_I^{(k)}\big(x_I+\xi_I, \bar{x}_I+\bar{\xi}_I, u_I\big)|. \qquad (3.23)$$

Since the regions $\mathcal{D}_I$, $\mathcal{D}_{\xi_I}$ and $\mathcal{D}_{\bar{\xi}_I}$ are compact sets, the supremums in (3.22) and (3.23) are finite. In addition, note that the bound $\bar{\chi}_I^{(k)}(t)$ in (3.21) depends on t because of the bounding function $\bar{\eta}_I^{(k)}$.

The first term on the right-hand side of (3.19) can be bounded by the output of a suitably selected filter $\bar{H}_p(s)$ with impulse response $\bar{h}_p(t)$ such that $|h_p(t)| \leq \bar{h}_p(t)$ for all $t > 0$. Details on calculating a suitable transfer function $\bar{H}_p(s)$ will be given in the sequel in Section 3.2.2. Therefore, the first term on the right-hand side of (3.19) can be bounded as follows:

$$\int_0^t |h_p(t-\tau)|\bar{\chi}_I^{(k)}(\tau)\,\mathrm{d}\tau \leq \bar{H}_p(s)\left[\bar{\chi}_I^{(k)}(t)\right].$$

Finally using (3.19), a suitable detection threshold $\bar{r}_I^{(k)}(t)$ is given by

$$\bar{r}_I^{(k)}(t) = \int_0^t \bar{h}_p(t-\tau)\bar{\chi}_I^{(k)}(\tau)\,\mathrm{d}\tau + \bar{\epsilon}_{\xi_I}^{(k)} + \bar{\epsilon}_{\zeta_I}^{(k)}, \qquad (3.24)$$

Christodoulos Keliris

and can be implemented by

$$\bar{r}_I^{(k)}(t) = \bar{H}_p(s)\left[\bar{\chi}_I^{(k)}(t)\right] + \bar{\epsilon}_{\xi_I}^{(k)} + \bar{\epsilon}_{\zeta_I}^{(k)}. \tag{3.25}$$

A practical issue that requires consideration is the derivation of the bound $\bar{\chi}_I^{(k)}(t)$ given in (3.21). Specifically, the derivation of $\bar{\chi}_I^{(k)}(t)$ requires the bounds $\overline{\Delta f}_I^{(k)}$ and $\overline{\Delta g}_I^{(k)}$ on $\Delta f_I^{(k)}(t)$ and $\Delta g_I^{(k)}(t)$ respectively. One approach for deriving the bound $\overline{\Delta f}_I^{(k)}$ in (3.22) is to consider a local Lipschitz assumption:

$$|f_I^{(k)}(x_I, u_I) - f_I^{(k)}(x_I + \xi_I, u_I)| \leq L_{f_I^{(k)}}|\xi_I| \tag{3.26}$$

where $L_{f_I^{(k)}}$ is the Lipschitz constant for the function $f_I^{(k)}(x_I, u_I)$ with respect to x_I in the region $\mathcal{D}_{x_I}$. Therefore, if we have a bound ξ_I^M on the measurement noise, i.e. $|\xi_I(t)| \leq \xi_I^M \quad \forall t > 0$, then we can obtain a bound on $\Delta f_I^{(k)}(t)$. A similar approach can be followed for the interconnection functions $\Delta g_I^{(k)}(t)$.

Filtering can also be proved beneficial for dampening the mismatch function $\Delta f_I(t) + \Delta g_I(t)$ which results due to the measurement noise and therefore further enhance fault detectability. Among the various filters one can select, some may lead to less conservative detection thresholds. Therefore, a significantly less conservative detection threshold without the need for the Lipschitz constants can be obtained by observing that the residual (3.15) can be written as

$$r_I(t) = H_p(s)\left[\eta_I\big(x_I(t), \bar{x}_I(t), u_I(t), t\big)\right] + H_p(s)\left[\Delta f_I(t) + \Delta g_I(t)\right] + \epsilon_{\xi_I}(t) + \epsilon_{\zeta_I}(t) \tag{3.27}$$

and by making the following assumption:

***Assumption* 3.3.** *The filtered function mismatch term $\epsilon_{\Delta_I}(t) \triangleq H_p(s)\left[\Delta f_I(t) + \Delta g_I(t)\right]$ is uniformly bounded as follows:*

$$|\epsilon_{\Delta_I}^{(k)}(t)| \leq \bar{\epsilon}_{\Delta_I}^{(k)} \quad k = 1, 2, \ldots, n_I, \tag{3.28}$$

where $\bar{\epsilon}_{\Delta_I}^{(k)}$ is a known bounding constant.

Assumption 3.3 is based on the fact that filtering dampens the effect of noise and disturbances present in the function mismatch term $\Delta f_I(t) + \Delta g_I(t)$. A suitable selection of $\bar{\epsilon}_{\Delta_I}^{(k)}$ can be made through the use of simulations (i.e. Monte Carlo methods) by filtering the function mismatch term $\Delta f_I(t) + \Delta g_I(t)$ using the known nominal function dynamics and the available noise characteristics (recall that the measurement noise is assumed to take values in a compact set).

Christodoulos Keliris

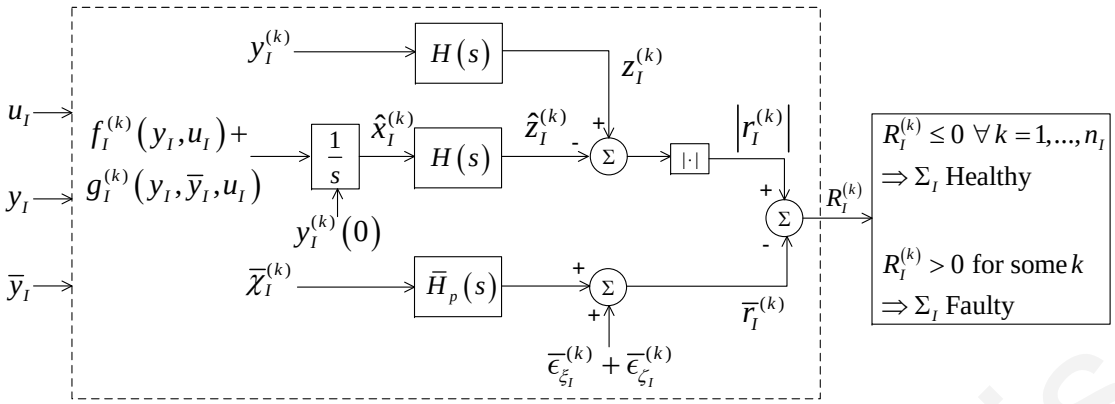


Figure 3.2: Local Filtered Fault Detection Scheme

Therefore, the detection threshold becomes

$$\bar{r}_I^{(k)}(t) = \bar{H}_p(s)\left[\bar{\eta}_I^{(k)}\big(y_I(t),\bar{y}_I(t),u_I(t)\big)\right] + \bar{\epsilon}_{\Delta_I}^{(k)} + \bar{\epsilon}_{\xi_I}^{(k)} + \bar{\epsilon}_{\zeta_I}^{(k)}. \tag{3.29}$$

Figure 3.2 illustrates the implementation of the local filtered fault detection scheme for the I-th subsystem resulting from equations (3.9), (3.11), (3.12), (3.14) and (3.25).

The findings of the preceding analysis are summarized in the following lemma:

Lemma 3.1. *Consider a distributed system constituted by N subsystems Σ_I given by (3.1), (3.2) with the fault detection scheme described by (3.9), (3.11) and (3.12). Then for any $k = 1, 2, \ldots, n_I$, the residual signal given by (3.14) satisfies*

$$|r_I^{(k)}(t)| \le \bar{r}_I^{(k)}(t), \quad \forall t \in [0, T_0).$$

Lemma 3.1 means that, in the absence of any faults the absolute value of the residual signal $r_I^{(k)}(t)$ is always bounded by the detection threshold $\bar{r}_I^{(k)}(t),\ k = 1, 2, \ldots, n_I,\ \forall\, t \ge 0$ given by (3.25) or (3.29) thus guaranteeing that there will be no false alarms.

The same filtering approach can be applied in the ideal case of no noise or disturbances (i.e. $\epsilon_{\xi_I}(t) = \epsilon_{\zeta_I}(t) = 0$) in order to detect faults in a certain frequency range. In this case a first order filter $H(s) = \frac{d_0 s}{s+c_0}$, which implies that $H_p(s) = \frac{d_0}{s+c_0}$, can also be used. However, in the presence of measurement noise, the strictly proper requirement for $H(s)$ is important. It is interesting to examine the case where $H(s)$ is not strictly proper by considering the simple case of a first-order transfer function $H(s) = \frac{d_0 s}{s+c_0}$, which implies that $H_p(s) = \frac{d_0}{s+c_0}$. It is easily seen that in this case the impulse response $h_p(t) = d_0 e^{-c_0 t}$ is non-negative for $d_0 > 0$. First, let's consider the case where there is no measurement noise or disturbances affecting the system, i.e. $\xi_I(t) = 0$, $\zeta_I(t) = 0$ for all $t > 0$, which implies that $\Delta f_I(t) = 0$, $\Delta g_I(t) = 0$ and the total uncertainty term (3.16) is simplified to

Christodoulos Keliris

$\chi_I(t) = \eta_I\big(x_I(t), \bar{x}_I(t), u_I(t), t\big)$. Therefore, the residual from (3.15) is now given by:

$$r_I^{(k)}(t) = \frac{d_0}{s + c_0}\big[\eta_I^{(k)}\big(x_I(t), \bar{x}_I(t), u_I(t), t\big)\big]$$

and a suitable detection threshold is:

$$\bar{r}_I^{(k)}(t) = \frac{d_0}{s + c_0}\big[\bar{\eta}_I^{(k)}\big(x_I(t), \bar{x}_I(t), u_I(t)\big)\big]. \tag{3.30}$$

The detection threshold in (3.30) is sufficient when no noise or disturbances affect the system. However, in the presence of noise or disturbances, the use of (3.30) may easily lead to false alarms, since the effects of noise are not taken into consideration by the detection threshold. This will be further demonstrated in the simulation results. Now, consider the case where the presence of noise and disturbances is acknowledged and their effects are counteracted by using appropriate bounds on their magnitude in order to avoid potential false alarms. More specifically, let the measurement noise and disturbance be bounded by positive constants, i.e.: $|\xi_I^{(k)}(t)| \leq \xi_{I,b}^{(k)}$, $|\zeta_I^{(k)}(t)| \leq \zeta_{I,b}^{(k)}$. In this case, the signal $z_I(t) = H(s)\big[y_I(t)\big]$ becomes:

$$\begin{aligned} z_I(t) &= \frac{s}{s + c_0}\big[d_0 x_I(t)\big] + \frac{s}{s + c_0}\big[d_0 \xi_I(t)\big] \\ &= \frac{d_0}{s + c_0}\big[f_I\big(x_I(t), u_I(t)\big) + g_I\big(x_I(t), \bar{x}_I(t), u_I(t)\big) + \zeta_I(t) + \eta_I\big(x_I(t), \bar{x}_I(t), u_I(t), t\big)\big] \\ &\quad + d_0\xi_I(t) - \frac{d_0}{s + c_0}\big[c_0\xi_I(t)\big]. \end{aligned}$$

Then, the residual $r_I(t) = z_I(t) - \hat{z}_I(t)$, where

$$\hat{z}_I(t) = \frac{d_0}{s + c_0}\big[f_I\big(y_I(t), u_I(t)\big) + g_I\big(y_I(t), \bar{y}_I(t), u_I\big)\big],$$

prior to the fault becomes

$$\begin{aligned} r_I^{(k)}(t) =& \frac{d_0}{s + c_0}\big[\Delta f_I^{(k)}(t) + \Delta g_I^{(k)}(t) - c_0\xi_I^{(k)}(t) + \eta_I^{(k)}\big(x_I(t), \bar{x}_I(t), u_I(t), t\big) + \zeta_I^{(k)}(t)\big] \\ &+ d_0\xi_I^{(k)}(t). \end{aligned}$$

In this case, the detection threshold is given by

$$\bar{r}_I^{(k)}(t) = \frac{d_0}{s + c_0}\big[\bar{\chi}_I^{(k)}(t) + \zeta_{I,b}^{(k)} + c_0\xi_{I,b}^{(k)}\big] + d_0\xi_{I,b}^{(k)}. \tag{3.31}$$

It is important to note that the derived detection threshold (3.31) contains additional terms regarding the bounds of the measurement noise and of the disturbance. Therefore, a first order filter makes the detection scheme more conservative, which may lead to missed detection. This is further illustrated in the simulation results (Section 3.5).

Christodoulos Keliris

3.2.2 Selection of filter $\bar{H}_p(s)$

In this part, two methods for selecting a suitable transfer function $\bar{H}_p(s)$ with impulse response $\bar{h}_p(t)$ such that $|h_p(t)| \leq \bar{h}_p(t)$ for all $t \geq 0$ are proposed.

In general though, note that if the impulse response $h_p(t)$ is non-negative, i.e. $h_p(t) \geq 0$, for all $t \geq 0$, then the calculation of $\bar{H}_p(s)$ can be omitted. In this case $H_p(s)$ can be used instead of $\bar{H}_p(s)$ in (3.25), as it can easily be seen from (3.19) where $|h_p(t-\tau)| = h_p(t-\tau)$. Necessary and sufficient conditions for non-negative impulse response for a specific class of filters are given in [59].

- *First method.*

The first method relies on the following lemma, which describes a methodology for finding $\bar{H}_p(s)$. For notational convenience, for any $m \times n$ matrix A we define $|A|_{\mathcal{E}}$ as the matrix whose elements correspond to the modulus of the element $a_{i,j}$, $i = 1, \ldots, m$ and $j = 1, \ldots, n$ of the matrix A; i.e.,

$$|A|_{\mathcal{E}} \triangleq \begin{bmatrix} |a_{1,1}| & \ldots & |a_{1,n}| \\ \vdots & \ddots & \vdots \\ |a_{m,1}| & \ldots & |a_{m,n}| \end{bmatrix}.$$

Lemma 3.2. *Let $w(t) = Ce^{At}B$ be the impulse response of a strictly proper SISO transfer function $W(s)$ with state space representation (A, B, C). Then, for any signal $v(t) \geq 0$, the following inequality holds for all $t \geq 0$:*

$$\int_0^t |w(t-\tau)|v(\tau)\,\mathrm{d}\tau \leq \overline{W}(s)\,[v(t)]\,,$$

where $\overline{W}(s)$ is given by

$$\overline{W}(s) \triangleq |CT|_{\mathcal{E}}\,(sI - \boldsymbol{Re}[J])^{-1}\left|T^{-1}B\right|_{\mathcal{E}} \tag{3.32}$$

and $J = T^{-1}AT$ is the Jordan form of the matrix A.

Proof. Using Jordan Theorem, A can be written as $A = TJT^{-1}$ where

$$J = diag\big[J_1(\lambda_1), J_2(\lambda_2), \cdots, J_l(\lambda_l)\big]$$

and $J_k(\lambda_k)$ is the Jordan block associated with eigenvalue $\lambda_k \triangleq \sigma_k + \jmath\omega_k, k = 1, \cdots, l$. Using $e^{At} = Te^{Jt}T^{-1}$ [7], we have

$$\begin{aligned} |w(t)| &= |CTe^{Jt}T^{-1}B| \\ &\leq |CT|_{\mathcal{E}}\left|e^{Jt}\right|_{\mathcal{E}}\left|T^{-1}B\right|_{\mathcal{E}}. \end{aligned} \tag{3.33}$$

Christodoulos Keliris

Now, by splitting J in its real and imaginary parts, i.e. $J = \text{Re}[J] + \jmath\text{Im}[J]$, (3.33) satisfies

$$\begin{aligned} |w(t)| &\leq |CT|_{\mathcal{E}} \left|e^{\text{Re}[J]t}\right|_{\mathcal{E}} \left|e^{\jmath\text{Im}[J]t}\right|_{\mathcal{E}} \left|T^{-1}B\right|_{\mathcal{E}} \\ &= |CT|_{\mathcal{E}} \left|e^{\text{Re}[J]t}\right|_{\mathcal{E}} \left|T^{-1}B\right|_{\mathcal{E}}. \end{aligned} \tag{3.34}$$

The last equation results from the fact that $\left|e^{\jmath\text{Im}[J]t}\right|_{\mathcal{E}} = I$ since $\text{Im}[J]$ is a diagonal matrix. By using the property of the exponential matrix [7]

$$e^{\text{Re}[J]t} = diag\left[e^{J_1(\sigma_1)t}, \cdots, e^{J_l(\sigma_l)t}\right],$$

where the elements of the exponential form of the k-th Jordan block $e^{J_k(\sigma_k)t}, k = 1, \cdots, l$ with size $\nu_k \times \nu_k$ are given by [15]

$$e^{J_k(\sigma_k)t}_{i,j} = \begin{cases} e^{\sigma_k t}\frac{t^{j-1}}{(j-1)!} & \text{for } 1 \leq i \leq j \leq \nu_k \\ 0 & \text{otherwise.} \end{cases} \tag{3.35}$$

From (3.35) it is easily seen that all the elements of $e^{J_k(\sigma_k)t}, k = 1, \cdots, l$ are non-negative. Therefore $\left|e^{\text{Re}[J]t}\right|_{\mathcal{E}} = e^{\text{Re}[J]t}$ and finally (3.34) becomes $|w(t)| \leq \bar{w}(t)$ where

$$\bar{w}(t) \triangleq |CT|_{\mathcal{E}}\, e^{\text{Re}[J]t} \left|T^{-1}B\right|_{\mathcal{E}}.$$

Since $v(t) \geq 0$, therefore the following inequality is satisfied for all $t \geq 0$

$$\begin{aligned} \int_0^t |w(t-\tau)|v(\tau)\,\mathrm{d}\tau &\leq \int_0^t \bar{w}(t-\tau)v(\tau)\,\mathrm{d}\tau \\ &= \overline{W}(s)\left[v(t)\right], \end{aligned}$$

where the transfer function $\overline{W}(s)$ can be easily implemented as

$$\overline{W}(s) \triangleq |CT|_{\mathcal{E}} \left(sI - \text{Re}[J]\right)^{-1} \left|T^{-1}B\right|_{\mathcal{E}}.$$

□

Therefore, by using Lemma 3.2, a bound $\bar{h}_p(t)$ such that $|h_p(t)| \leq \bar{h}_p(t)$ and its respective transfer function $\bar{H}_p(s)$ can be obtained.

- *Second method.*

The second method is by using the following well-known result (see, for instance [21]).

Lemma 3.3. *The impulse response $h_p(t)$ of a strictly proper and asymptotically stable transfer function $H_p(s)$ decays exponentially; i.e., $|h_p(t)| \leq \kappa e^{-\upsilon t}$ for some $\kappa > 0$, $\upsilon > 0$, for all $t \geq 0$.*

By using Lemma 3.3, a suitable impulse response $\bar{h}_p(t)$ such that $|h_p(t)| \leq \bar{h}_p(t)$ for all $t \geq 0$ is given by $\bar{h}_p(t) = \kappa e^{-\upsilon t}$ and can be implemented using linear filtering techniques as $\bar{H}_p(s) = \frac{\kappa}{s+\upsilon}$.

Christodoulos Keliris

3.3 Fault Detectability Analysis

The design and analysis of the fault detection scheme in the previous two sections were based on the derivation of suitable thresholds $\bar{r}_I^{(k)}(t)$ such that in the absence of any fault, the residual signals $r_I^{(k)}(t)$ are bounded by $\bar{r}_I^{(k)}(t)$. An important related question is what class of faults can be detected. This is referred to as *fault detectability analysis*. In this section, fault detectability conditions for the aforementioned fault detection scheme are derived. The fault detectability analysis constitutes a theoretical result that characterizes quantitatively the class of faults detectable by the proposed scheme.

Theorem 3.1. *Consider the nonlinear system (3.1), (3.2) with the distributed fault detection scheme described in (3.9), (3.11), (3.12), (3.14) and (3.24) in the general case of $H(s)$ given by (3.5). A fault in the I-th subsystem initiated at $t = T_0$ is detectable if the fault function $\phi_I^{(k)}(x, u_I)$ satisfies the following inequality for some $k = 1, 2, \ldots, n_I$:*

$$\left| \int_{T_0}^{t} h_p(t-\tau) \left(1 - e^{-b_I(\tau - T_0)}\right) \phi_I^{(k)}\big(x(\tau), u_I(\tau)\big)\, \mathrm{d}\tau \right| > 2\bar{r}_I^{(k)}(t). \tag{3.36}$$

Proof. In the presence of a fault that occurs at $t = T_0$, equation (3.15) becomes:

$$r_I^{(k)}(t) = H_p(s) \left[\chi_I^{(k)}(t) + \beta_I(t - T_0)\phi_I^{(k)}\big(x(t), u_I(t)\big)\right] + \epsilon_{\xi_I}^{(k)}(t) + \epsilon_{\zeta_I}^{(k)}(t).$$

By using the triangle inequality, for $t > T_0$, the residual $r_I^{(k)}(t)$ satisfies:

$$\begin{aligned}
|r_I^{(k)}(t)| \geq & - |H_p(s) \left[\chi_I^{(k)}(t)\right]| - |\epsilon_{\xi_I}^{(k)}(t)| - |\epsilon_{\zeta_I}^{(k)}(t)| \\
& + |H_p(s) \left[\beta_I(t - T_0)\phi_I^{(k)}\big(x(t), u_I(t)\big)\right]| \\
\geq & - \int_0^t |h_p(t-\tau)||\chi_I^{(k)}(\tau)|\, \mathrm{d}\tau - |\epsilon_{\xi_I}^{(k)}(t)| - |\epsilon_{\zeta_I}^{(k)}(t)| \\
& + |H_p(s) \left[\beta_I(t - T_0)\phi_I^{(k)}\big(x(t), u_I(t)\big)\right]| \\
\geq & - \int_0^t \bar{h}_p(t-\tau)\bar{\chi}_I^{(k)}(\tau)\, \mathrm{d}\tau - \bar{\epsilon}_{\xi_I}^{(k)} - \bar{\epsilon}_{\zeta_I}^{(k)} + |H_p(s) \left[\beta_I(t - T_0)\phi_I^{(k)}\big(x(t), u_I(t)\big)\right]| \\
\geq & - \bar{r}_I^{(k)}(t) + |H_p(s) \left[\beta_I(t - T_0)\phi_I^{(k)}(x, u_I)\right]|.
\end{aligned}$$

For fault detection, the inequality $|r_I^{(k)}(t)| > \bar{r}_I^{(k)}(t)$ must hold for some $k = 1, 2, \ldots, n_I$, so the final fault detectability condition is obtained:

$$|H_p(s) \left[\beta_I(t - T_0)\phi_I^{(k)}\big(x(t), u_I(t)\big)\right]| > 2\bar{r}_I^{(k)}(t).$$

This can be rewritten based on (3.3) in the integral form (3.36) of the Theorem. □

Christodoulos Keliris

Although Theorem 3.1 is based on threshold (3.24), it can be readily shown that the same result holds in the case where threshold (3.29) is used. Clearly, the fault functions $\phi_I(x, u_I)$ are typically unknown and therefore this condition cannot be checked apriori. However, it provides useful intuition about the type of faults that are detectable. The detectability condition given in Theorem 3.1 is a sufficient condition, but not a necessary one and hence, the class of detectable faults can be significantly larger. The use of filtering is of crucial importance in order to derive tight detection thresholds that guarantee no false alarms. As it can be seen in the detectability conditions given by (3.36) the detection of the fault depends on the filtered fault function ϕ_I and as a result, the selection of the filter is very important. Since the fault function is usually comprised of lower frequency components, it is not affected that much by low-pass filtering in comparison to the measurement noise which is usually of higher frequency. In addition, filtering allows the derivation of tighter detection thresholds and, as a result, the fault detectability condition can be met more easily. Obviously, some filter selections may lead to less conservative thresholds than others.

The detectability properties of the proposed filtering approach are further investigated by considering a specific case for the filter $H_p(s)$:

$$H_p(s) = \frac{\alpha^p}{(s+\alpha)^p}. \tag{3.37}$$

This type of filter is well suited for gaining further intuition since it contains two parameters p and α that denote the order of the filter and the pole location, respectively. More specifically, the order p of the filter regulates the damping effect of the high frequency noise, whereas the value α of the filter determines the cutoff frequency at which the damping begins. In general, more selective filter implementations can be made (i.e. Butterworth filters) which may have some implications in the filters required for the detection threshold implementation (due to the fact that the impulse response may not be always positive). But, the particular filter $H_p(s)$ given by (3.37) is perfectly suited for the investigation of the analytical properties of the filtering scheme.

In order to conduct this fault detectability analysis, we simplify Assumption 2 by considering a constant bounding condition.

***Assumption* 3.4.** *The modeling uncertainty $\eta_I^{(k)}$ in each subsystem is an unstructured and possibly unknown nonlinear function of x_I, $\bar{x}_I$, u_I and t but uniformly bounded by a known positive scalar $\bar{\eta}_I^{(k)}$, i.e.,*

$$|\eta_I^{(k)}(x_I, \bar{x}_I, u_I, t)| \leq \bar{\eta}_I^{(k)}, \quad k = 1, 2, \ldots, n_I \tag{3.38}$$

Christodoulos Keliris

for all $t \geq 0$ *and for all* $(x_I, \bar{x}_I, u_I) \in \mathcal{D}_I$*, where* $\bar{\eta}_I^{(k)} \geq 0$ *is a known bounding scalar in some region of interest* $\mathcal{D}_I = \mathcal{D}_{x_I} \times \mathcal{D}_{\bar{x}_I} \times \mathcal{D}_{u_I} \subset \mathbb{R}^{n_I} \times \mathbb{R}^{\bar{n}_I} \times \mathbb{R}^{m_I}$.

By using the Lipschitz assumption stated in (3.26), along with the known bound of the measurement uncertainty $|\xi_I|$ and the constant bound on the modeling uncertainty $\bar{\eta}_I^{(k)}$, as stated in Assumption 3.4, the bound of the total uncertainty term $\bar{\chi}_I^{(k)}(t)$ takes a constant value $\bar{\chi}_I^{(k)}$.

Although we use (3.25) for the detection threshold, the adaptation of the subsequent analysis in the case where the threshold is given by (3.29) is straightforward by simply replacing $\bar{\chi}_I^{(k)}$ with $\bar{\eta}_I^{(k)}$ and adding the term $\bar{\epsilon}_{\Delta_I}^{(k)}$ along the terms $\bar{\epsilon}_{\xi_I}^{(k)} + \bar{\epsilon}_{\zeta_I}^{(k)}$ in the subsequent analysis.

Next, some of the expressions given in the previous section will be rewritten according to the specific case of $H_p(s)$ given by (3.37). In this case, the inverse Laplace transform of $H_p(s)$ is given by:

$$h_p(t) = \frac{1}{(p-1)!} t^{p-1} e^{-\alpha t} \alpha^p. \tag{3.39}$$

For this case of the filter, the residual $r_I^{(k)}(t)$ from (3.15) becomes

$$r_I^{(k)}(t) = \frac{\alpha^p}{(s+\alpha)^p} \left[\chi_I^{(k)}(t) \right] + \epsilon_{\xi_I}^{(k)}(t) + \epsilon_{\zeta_I}^{(k)}(t) \tag{3.40}$$

and the detection threshold $\bar{r}_I^{(k)}(t)$ from (3.25) (since $h_p(t)$ is non-negative) becomes

$$\bar{r}_I^{(k)}(t) = \frac{\alpha^p}{(s+\alpha)^p} \left[\bar{\chi}_I^{(k)} \right] + \bar{\epsilon}_{\xi_I}^{(k)} + \bar{\epsilon}_{\zeta_I}^{(k)}. \tag{3.41}$$

Finally, the fault detectability condition in (3.36) becomes:

$$\frac{\alpha^p}{(p-1)!} \left| \int_{T_0}^{t} (t-\tau)^{p-1} e^{-\alpha(t-\tau)} \left(1 - e^{-b_I(\tau - T_0)}\right) \phi_I^{(k)}\big(x(\tau), u_I(\tau)\big) \, \mathrm{d}\tau \right| > 2\bar{r}_I^{(k)}(t). \tag{3.42}$$

Theorem 3.2. *Consider the nonlinear system (3.1), (3.2) with the distributed fault detection scheme described in (3.9), (3.11), (3.12), (3.14) and (3.41) in the special case of* $H_p(s)$ *given by (3.37). Suppose at least one component* $\phi_I^{(k)}(x, u_I)$ *of the fault vector* $\phi_I(x, u_I)$ *satisfies the condition*

$$|\phi_I^{(k)}(x(t'), u_I(t'))| \geq M, \quad \forall\, t' \in [T_0, t]\,, \tag{3.43}$$

for sufficiently large $t > T_0$ *and is continuous in the time interval* $t' \in [T_0, t]$*. If* $M > 2(\bar{\chi}_I^{(k)} + \bar{\epsilon}_{\xi_I}^{(k)} + \bar{\epsilon}_{\zeta_I}^{(k)})$*, then the fault will be detected, that is* $|r_I^{(k)}(t)| > \bar{r}_I^{(k)}(t)$.

Christodoulos Kallis

Proof. Since $\phi_I^{(k)}(x(t'), u_I(t'))$ is continuous for $t' \in [T_0, t]$ then (3.43) means that the fault function $\phi_I^{(k)}(x(t'), u_I(t'))$ does not change sign for $t' \in [T_0, t]$. Then, a sufficient condition for the fault detectability condition (3.42) to hold is:

$$\frac{\alpha^p}{(p-1)!} \int_{T_0}^{t} (t-\tau)^{p-1} e^{-\alpha(t-\tau)} \left(1 - e^{-b_I(\tau - T_0)}\right) M \, \mathrm{d}\tau > 2\bar{r}_I^{(k)}(t),$$

which implies that

$$\frac{M}{(p-1)!} \left(\int_{T_0}^{t} \alpha^p (t-\tau)^{p-1} e^{-\alpha(t-\tau)} \, \mathrm{d}\tau - \int_{T_0}^{t} \alpha^p e^{-b_I(\tau - T_0)} (t-\tau)^{p-1} e^{-\alpha(t-\tau)} \, \mathrm{d}\tau \right) > 2\bar{r}_I^{(k)}(t).$$

The previous inequality can be written as:

$$\frac{M}{(p-1)!} q(t, T_0, \alpha) - 2\bar{r}_I^{(k)}(t) > 0 \tag{3.44}$$

where:

$$q(t, T_0, \alpha) \triangleq q_1(t, T_0, \alpha) - q_2(t, T_0, \alpha) \tag{3.45}$$

$$q_1(t, T_0, \alpha) \triangleq \int_{T_0}^{t} \alpha^p (t-\tau)^{p-1} e^{-\alpha(t-\tau)} \, \mathrm{d}\tau \tag{3.46}$$

$$\begin{aligned} q_2(t, T_0, \alpha) &\triangleq \int_{T_0}^{t} \alpha^p (t-\tau)^{p-1} e^{-\alpha(t-\tau)} e^{-b_I(\tau - T_0)} \, \mathrm{d}\tau \\ &= e^{-b_I(t-T_0)} \int_{T_0}^{t} \alpha^p (t-\tau)^{p-1} e^{-(\alpha - b_I)(t-\tau)} \, \mathrm{d}\tau. \end{aligned} \tag{3.47}$$

Note that by definition $q_1(t, T_0, \alpha)$ and $q_2(t, T_0, \alpha)$ are always positive for any $t > T_0$, $\alpha > 0$, $p > 0$ and $b_I > 0$. By using the change of variables $w = \alpha(t - \tau)$, the function $q_1(t, T_0, \alpha)$ is transformed into:

$$q_1(t, T_0, \alpha) = \int_{0}^{\alpha(t-T_0)} w^{p-1} e^{-w} \, \mathrm{d}w. \tag{3.48}$$

In the sequel we will utilize the definition of Gamma functions [63]. Specifically, the Gamma function $\Gamma(p)$ and the upper incomplete Gamma function $\Gamma(p, z)$ are defined as follows:

$$\Gamma(p) \triangleq \int_{0}^{\infty} w^{p-1} e^{-w} \, \mathrm{d}w \tag{3.49}$$

$$\Gamma\big(p, z\big) \triangleq \int_{z}^{\infty} w^{p-1} e^{-w} \, \mathrm{d}w. \tag{3.50}$$

Then, the lower incomplete Gamma function $\gamma\big(p, z\big)$ is defined as:

$$\gamma\big(p, z\big) \triangleq \int_{0}^{z} w^{p-1} e^{-w} \, \mathrm{d}w = \Gamma(p) - \Gamma\big(p, z\big). \tag{3.51}$$

Note that if p is a positive integer (such as in our case where it denotes the order of the filter), then $\Gamma(p) = (p-1)!$.

Christodoulos Kellinis

According to these definitions, $q_1(t, T_0, \alpha)$ from (3.48) is written as:

$$q_1(t, T_0, \alpha) = \gamma\big(p, \alpha(t - T_0)\big). \tag{3.52}$$

Following a similar procedure and after some calculations, the term $q_2(t, T_0, \alpha)$ given in equation (3.47) can be written as:

$$q_2(t, T_0, \alpha) = \begin{cases} \frac{\alpha^p}{p}(t - T_0)^p e^{-\alpha(t-T_0)} & \text{if } \alpha = b_I \\ \frac{\alpha^p e^{-b_I(t-T_0)}}{(a-b_I)^p}\gamma\big(p, (\alpha - b_I)(t - T_0)\big) & \text{else.} \end{cases} \tag{3.53}$$

By using (3.48) and (3.49), the following limits are derived:

$$\begin{aligned} &\lim_{t\to T_0} q_1(t, T_0, \alpha) = 0 \\ &\lim_{t\to\infty} q_1(t, T_0, \alpha) = \Gamma(p) = (p-1)!. \end{aligned} \tag{3.54}$$

The respective limits for $q_2(t, T_0, \alpha)$ from (3.53) are given in both cases $\alpha = b_I$ and $\alpha \neq b_I$ by:

$$\begin{aligned} &\lim_{t\to T_0} q_2(t, T_0, \alpha) = 0 \\ &\lim_{t\to\infty} q_2(t, T_0, \alpha) = 0. \end{aligned} \tag{3.55}$$

Next, it will be proved that the quantity $q(t, T_0, \alpha)$ is monotonically increasing. To compute $\frac{\partial}{\partial t}q(t, T_0, \alpha)$, we first consider the partial derivative of $\gamma\big(p, \alpha(t - T_0)\big)$, which satisfies

$$\begin{aligned} \frac{\partial}{\partial t}\Big(\gamma\big(p, \alpha(t - T_0)\big)\Big) &= \frac{\partial}{\partial t}\Big(\int_0^{\alpha(t-T_0)} w^{p-1}e^{-w}\,\mathrm{d}w\Big) \\ &= \alpha\big(w^{p-1}e^{-w}\big)\Big|_{w=\alpha(t-T_0)} \\ &= \alpha^p(t - T_0)^{p-1}e^{-\alpha(t-T_0)}. \end{aligned} \tag{3.56}$$

First consider the case $\alpha = b_I$. Using equations (3.52), (3.53) and (3.56) the partial derivative of $q(t, T_0, \alpha)$ with respect to t is:

$$\begin{aligned} \frac{\partial q(t, T_0, \alpha)}{\partial t} &= \frac{\partial q_1(t, T_0, \alpha)}{\partial t} - \frac{\partial q_2(t, T_0, \alpha)}{\partial t} \\ &= \frac{\partial}{\partial t}\Big(\gamma\big(p, \alpha(t - T_0)\big)\Big) - \frac{\partial}{\partial t}\Big(\frac{\alpha^p}{p}(t - T_0)^p e^{-a(t-T_0)}\Big) \\ &= \alpha^p(t - T_0)^{p-1}e^{-\alpha(t-T_0)} - \alpha^p(t - T_0)^{p-1}e^{-\alpha(t-T_0)} + \frac{\alpha^{p+1}}{p}(t - T_0)^p e^{-\alpha(t-T_0)} \\ &= \frac{\alpha^{p+1}}{p}(t - T_0)^p e^{-\alpha(t-T_0)}, \end{aligned}$$

which is positive for all $t > T_0, \alpha > 0, p > 0$.

Christodoulos Keliris

Now consider the case $\alpha \neq b_I$. The partial derivative of $q(t, T_0, \alpha)$ with respect to t is:

$$\begin{aligned}
\frac{\partial q(t,T_0,\alpha)}{\partial t} &= \frac{\partial}{\partial t}\Big(\gamma\big(p, \alpha(t-T_0)\big)\Big) - \frac{\partial}{\partial t}\left(\frac{\alpha^p e^{-b_I(t-T_0)}}{(a-b_I)^p}\gamma\big(p,(\alpha-b_I)(t-T_0)\big)\right) \\
&= \alpha^p (t-T_0)^{p-1} e^{-\alpha(t-T_0)} + b_I q_2(t,T_0,\alpha) \\
&\quad - \frac{\alpha^p}{(a-b_I)^p} e^{-b_I(t-T_0)} \frac{\partial}{\partial t}\Big(\gamma\big(p,(\alpha-b_I)(t-T_0)\big)\Big) \\
&= \alpha^p (t-T_0)^{p-1} e^{-\alpha(t-T_0)} + b_I q_2(t,T_0,\alpha) \\
&\quad - \alpha^p e^{-b_I(t-T_0)} (t-T_0)^{p-1} e^{-(\alpha-b_I)(t-T_0)} \\
&= b_I q_2(t,T_0,\alpha),
\end{aligned}$$

which is again positive for all $t > T_0, \alpha > 0, p > 0, b_I > 0$.

Therefore $q(t, T_0, \alpha)$ is monotonically increasing for any positive α, p, b_I, and its limits are derived by combining equations (3.54) and (3.55):

$$\lim_{t\to T_0} q(t,T_0,\alpha) = 0 \tag{3.57}$$

$$\lim_{t\to\infty} q(t,T_0,\alpha) = \Gamma(p) = (p-1)!. \tag{3.58}$$

The detection threshold given in (3.41) can be written as

$$\bar{r}_I^{(k)}(t) = \int_0^t \frac{1}{(p-1)!}(t-\tau)^{p-1} e^{-\alpha(t-\tau)} \alpha^p \bar{\chi}_I^{(k)} \,\mathrm{d}\tau + \bar{\epsilon}_{\xi_I}^{(k)} + \bar{\epsilon}_{\zeta_I}^{(k)},$$

and since the bound of the total uncertainty term $\bar{\chi}_I^{(k)}$ has a constant value (as explained earlier), the previous equation can be written as

$$\bar{r}_I^{(k)}(t) = \frac{1}{(p-1)!}\bar{\chi}_I^{(k)} \gamma\big(p, \alpha t\big) + \bar{\epsilon}_{\xi_I}^{(k)} + \bar{\epsilon}_{\zeta_I}^{(k)}. \tag{3.59}$$

Note that $\bar{r}_I^{(k)}(t)$ is a monotonically increasing function since

$$\frac{\partial \bar{r}_I^{(k)}(t)}{\partial t} = \frac{1}{(p-1)!}\bar{\chi}_I^{(k)} \alpha^p t^{p-1} e^{-\alpha t} > 0,$$

with limits:

$$\lim_{t\to 0} \bar{r}_I^{(k)}(t) = \bar{\epsilon}_{\xi_I}^{(k)} + \bar{\epsilon}_{\zeta_I}^{(k)} \tag{3.60}$$

$$\lim_{t\to\infty} \bar{r}_I^{(k)}(t) = \bar{\chi}_I^{(k)} + \bar{\epsilon}_{\xi_I}^{(k)} + \bar{\epsilon}_{\zeta_I}^{(k)}. \tag{3.61}$$

After sufficiently long time, the left side of (3.44) using (3.58) and (3.61) is:

$$\lim_{t\to\infty}\Big(\frac{M}{(p-1)!} q(t,T_0,\alpha) - 2\bar{r}_I^{(k)}(t)\Big) = M - 2(\bar{\chi}_I^{(k)} + \bar{\epsilon}_{\xi_I}^{(k)} + \bar{\epsilon}_{\zeta_I}^{(k)}).$$

Therefore, if $M > 2(\bar{\chi}_I^{(k)} + \bar{\epsilon}_{\xi_I}^{(k)} + \bar{\epsilon}_{\zeta_I}^{(k)})$ then (3.44) will be satisfied and the fault will be detected. □

Christodoulos Keliris

The aforementioned theorem is conceptually different from Theorem 3.1. More specifically, the detectability condition (3.36) of Theorem 3.1 allows the fault function $\phi_I^{(k)}$ to change sign. On the other hand, Theorem 3.2 states that if the fault function $\phi_I^{(k)}$ maintains the same sign over time and its magnitude is larger than $2(\bar{\chi}_I^{(k)} + \bar{\epsilon}_{\xi_I}^{(k)} + \bar{\epsilon}_{\zeta_I}^{(k)})$ for sufficiently long, then the fault is guaranteed to be detected. In the next section, an upper bound of the detection time is obtained in the case where a fault is detected according to Theorem 3.2. Moreover, we investigate the influence of the filter's order p and the pole location α on the upper bound of the detection time in order to derive some insight regarding the selection of p and α.

3.4 Detection Time Analysis

The detection time of a fault, that is, the time interval between the fault occurrence and its detection, plays a crucial role in fault diagnosis and it constitutes a form of performance criterion. When a fault is detected faster, then timely actions can be undertaken to avoid more serious or even disastrous consequences. It is worth noting that incipient faults are more difficult to detect, especially during their early stages, and as a result the detection time of an incipient fault is generally larger than that of an abrupt fault. The results are obtained for the general case of an incipient fault; concerning the dependence of the detection time on the filter's order p, only the abrupt fault case is addressed for the sake of simplicity.

Theorem 3.3. *Consider the nonlinear system (3.1), (3.2) with the distributed fault detection scheme described in (3.9), (3.11), (3.12), (3.14) and (3.41) in the special case of $H_p(s)$ given by (3.37). If at least one component $\phi_I^{(k)}(x, u_I)$ of the fault vector $\phi_I(x, u_I)$ satisfies the condition*

$$\left|\phi_I^{(k)}\big(x(t'), u_I(t')\big)\right| \geq M, \quad \forall\, t' \in [T_0, t] \tag{3.62}$$

where $M > 2(\bar{\chi}_I^{(k)} + \bar{\epsilon}_{\xi_I}^{(k)} + \bar{\epsilon}_{\zeta_I}^{(k)})$ for sufficiently large $t > T_0$ and is continuous in the time interval $t' \in [T_0, t]$ such that the fault can be detected according to Theorem 3.2, then:

(a) A sufficient condition for fault detectability is given by:

$$q(t, T_0, \alpha) > \frac{2(p-1)!}{M}(\bar{\chi}_I^{(k)} + \bar{\epsilon}_{\xi_I}^{(k)} + \bar{\epsilon}_{\zeta_I}^{(k)}). \tag{3.63}$$

(b) An upper bound on the detection time T_d of an incipient fault can be found by solving the equation:

$$q_1(T_d, T_0, \alpha) - q_2(T_d, T_0, \alpha) = \frac{2(p-1)!}{M}\bar{r}_I^{(k)}(T_d), \tag{3.64}$$

Christodoulos Kelinis

where q_1, q_2, $\bar{r}_I^{(k)}$ are given in (3.52), (3.53) and (3.59) respectively.

(c) The upper bound T_d decreases monotonically as the value of α increases.

(d) In the case of abrupt faults, the upper bound on the detection time T_d increases as the order p of the filter increases.

Proof. (a) By considering equations (3.44) and (3.61), a sufficient condition for fault detectability is given by (3.63).

(b) From (3.44) it easily follows that the upper bound on the detection time can be found by solving numerically the following equation:

$$q_1(T_d, T_0, \alpha) - q_2(T_d, T_0, \alpha) = \frac{2(p-1)!}{M}\bar{r}_I^{(k)}(T_d), \tag{3.65}$$

where q_1, q_2, $\bar{r}_I^{(k)}$ are given in (3.52), (3.53) and (3.59) respectively thus proving (3.65). Note that by using the sufficient condition (3.63) a more conservative upper bound on the detection time can be obtained from solving:

$$q_1(T_d, T_0, \alpha) - q_2(T_d, T_0, \alpha) = \frac{2(p-1)!}{M}(\bar{\chi}_I^{(k)} + \bar{\epsilon}_{\xi_I}^{(k)} + \bar{\epsilon}_{\zeta_I}^{(k)}). \tag{3.66}$$

(c) Next, we investigate the dependence of the bound on the detection time on the filter's pole α by considering the sufficient condition for fault detectability given by (3.63) and it will be proved that the quantity $q(T_d, T_0, \alpha)$ is monotonically increasing with respect to α. To compute $\frac{\partial}{\partial\alpha}q(T_d, T_0, \alpha)$, we first consider the partial derivative of $\gamma\big(p, \alpha(T_d - T_0)\big)$, which satisfies

$$\begin{aligned}\frac{\partial}{\partial\alpha}\Big(\gamma\big(p, \alpha(T_d - T_0)\big)\Big) &= \frac{\partial}{\partial\alpha}\bigg(\int_0^{\alpha(T_d-T_0)} w^{p-1}e^{-w}\,\mathrm{d}w\bigg)\\ &= (T_d - T_0)\big(w^{p-1}e^{-w}\big)\Big|_{w=\alpha(T_d-T_0)}\\ &= \alpha^{p-1}(T_d - T_0)^p e^{-\alpha(T_d-T_0)}.\end{aligned} \tag{3.67}$$

First consider the case $\alpha = b_I$. Using equations (3.52), (3.53) and (3.67), the partial derivative of $q(T_d, T_0, \alpha)$ with respect to α is:

$$\begin{aligned}\frac{\partial q(T_d, T_0, \alpha)}{\partial\alpha} &= \frac{\partial}{\partial\alpha}\Big(\gamma\big(p, \alpha(T_d - T_0)\big)\Big) - \frac{\partial}{\partial\alpha}\bigg(\frac{\alpha^p}{p}(T_d - T_0)^p e^{-\alpha(T_d-T_0)}\bigg)\\ &= \alpha^{p-1}(T_d - T_0)^p e^{-\alpha(T_d-T_0)} - p\alpha^{p-1}\frac{1}{p}(T_d - T_0)^p e^{-\alpha(T_d-T_0)}\\ &\quad + (T_d - T_0)e^{-\alpha(T_d-T_0)}\frac{\alpha^p}{p}(T_d - T_0)^p\\ &= \frac{\alpha^p}{p}(T_d - T_0)^{p+1}e^{-\alpha(T_d-T_0)},\end{aligned} \tag{3.68}$$

Christoforos Keliris

which is positive for all $T_d > T_0, \alpha > 0, p > 0$. So, in the case $\alpha = b_I$ the quantity $q(T_d, T_0, \alpha)$ is monotonically increasing with respect to α with limits:

$$\lim_{\alpha \to 0} q(T_d, T_0, \alpha) = 0$$
$$\lim_{\alpha \to \infty} q(T_d, T_0, \alpha) = \Gamma(p) = (p-1)!.$$

Now consider the case $\alpha \neq b_I$. The partial derivative of $q(T_d, T_0, \alpha)$ with respect to α is:

$$\frac{\partial q(T_d, T_0, \alpha)}{\partial \alpha} = \frac{\partial}{\partial \alpha}\Big(\gamma\big(p, \alpha(T_d - T_0)\big)\Big) - \frac{\partial}{\partial \alpha}\left(\frac{\alpha^p e^{-b_I(T_d - T_0)}}{(\alpha - b_I)^p}\right)\gamma\big(p, (\alpha - b_I)(T_d - T_0)\big)$$
$$- \frac{\alpha^p e^{-b_I(T_d - T_0)}}{(\alpha - b_I)^p}\frac{\partial}{\partial \alpha}\Big(\gamma\big(p, (\alpha - b_I)(T_d - T_0)\big)\Big).$$

Using (3.67) the previous equation becomes:

$$\frac{\partial q(T_d, T_0, \alpha)}{\partial \alpha} = \alpha^{p-1}(T_d - T_0)^p e^{-\alpha(T_d - T_0)} + \frac{\alpha^{p-1} b_I e^{-b_I(T_d - T_0)}}{(\alpha - b_I)^{p+1}} p\gamma\big(p, (\alpha - b_I)(T_d - T_0)\big)$$
$$- \frac{\alpha^p e^{-b_I(T_d - T_0)}}{\alpha - b_I}(T_d - T_0)^p e^{-(\alpha - b_I)(T_d - T_0)}. \tag{3.69}$$

Next, we use the recursive formula that is satisfied by the incomplete Gamma function [63]:

$$\gamma\big(p+1, z\big) = p\gamma\big(p, z\big) - z^p e^{-z}. \tag{3.70}$$

Using the aforementioned recursive formula with $z = (\alpha - b_I)(T_d - T_0)$ to substitute the relevant part in the second term of (3.69) and after some algebraic manipulation we get:

$$\frac{\partial q(T_d, T_0, \alpha)}{\partial \alpha} = \frac{\alpha^{p-1} b_I e^{-b_I(T_d - T_0)}}{(\alpha - b_I)^{p+1}}\gamma\big(p+1, (\alpha - b_I)(T_d - T_0)\big),$$

which is positive for all $T_d > T_0, \alpha > 0, b_I > 0, p > 0$, since it has a similar form to (3.53) in the case $\alpha \neq b_I$ (which is always positive as explained earlier). So, in the case $\alpha \neq b_I$ the quantity $q(T_d, T_0, \alpha)$ is monotonically increasing with respect to α with limits:

$$\lim_{\alpha \to 0} q(T_d, T_0, \alpha) = 0$$
$$\lim_{\alpha \to \infty} q(T_d, T_0, \alpha) = (1 - e^{-b_I(T_d - T_0)})\Gamma(p).$$

From the aforementioned analysis, it is clear that as α increases from 0 to ∞, $q(T_d, T_0, \alpha)$ increases monotonically from 0 to $\Gamma(p)$ in the case $\alpha = b_I$ or to $(1 - e^{-b_I(T_d - T_0)})\Gamma(p)$ in the case $\alpha \neq b_I$. Therefore, the upper bound on the detection time T_d decreases monotonically as α increases.

(d) In this part of the proof we investigate the effects that the order of the filter has on the upper bound of the detection time. For simplicity, we consider the sufficient detectability

Christoforos Kelins

condition given in (3.63) in the case of abrupt faults which corresponds to the limit $b_I \to \infty$, i.e. $q_2(t, T_0, \alpha) = 0$.

Consider the case in which two implementations of the proposed scheme are considered, the first one makes use of a p-th order filter and the second one a $p+1$ order filter. Consider also that a sufficiently large fault occurs and lasts enough so that both implementations are able to detect it, by using the same bounds $\bar{\epsilon}_{\xi_I}^{(k)}$ and $\bar{\epsilon}_{\zeta_I}^{(k)}$ so that $\bar{\psi}_I^{(k)} \triangleq \bar{\chi}_I^{(k)} + \bar{\epsilon}_{\xi_I}^{(k)} + \bar{\epsilon}_{\zeta_I}^{(k)}$ is the same for both implementations. For the first implementation that makes use of the p-th order filter, the upper bound on the detection time T_d can be found by solving (3.63) with $q_2(T_d, T_0, \alpha) = 0$, i.e.:

$$\gamma\big(p, \alpha(T_d - T_0)\big) = \frac{2(p-1)!}{M}\bar{\psi}_I^{(k)}. \tag{3.71}$$

For the second implementation making use of a $p+1$ order filter, we need to investigate its detectability condition at the time T_d. So, by moving all the terms of the detectability condition for the second implementation in the left hand side (replace p in (3.71) with $p+1$) we have:

$$\text{LHS} = \gamma\big(p+1, \alpha(T_d - T_0)\big) - \frac{2p!}{M}\bar{\psi}_I^{(k)}. \tag{3.72}$$

By using the recursive equation (3.70) with $z = \alpha(T_d - T_0)$, equation (3.72) becomes:

$$\begin{aligned}\text{LHS} &= p\gamma\big(p, \alpha(T_d - T_0)\big) - \alpha^p (T_d - T_0)^p e^{-\alpha(T_d - T_0)} - p\frac{2(p-1)!}{M}\bar{\psi}_I^{(k)} \\ &= p\left(\gamma\big(p, \alpha(T_d - T_0)\big) - \frac{2(p-1)!}{M}\bar{\psi}_I^{(k)}\right) - \alpha^p (T_d - T_0)^p e^{-\alpha(T_d - T_0)}.\end{aligned}$$

Then by using (3.71), the previous equation becomes:

$$\text{LHS} = -\alpha^p (T_d - T_0)^p e^{-\alpha(T_d - T_0)},$$

which is negative for all $T_d > T_0, \alpha > 0, p > 0$. Therefore, the detectability condition for the second implementation is not met at time T_d and so the upper bound on the detection time for the second implementation is larger than T_d. Since this holds for any value p, we conclude that increasing the order p of the filter results in increased upper bound for the detection time. □

Part (b) of the above theorem establishes the mathematical equation whose solution gives an upper bound on the detection time. At this point, we must stress that, although we refer to the solution of the equation as the upper bound of the detection time (because of the requirement (3.43)), there are cases where the solution is the actual detection time. For instance,

Christodoulos Kellinis

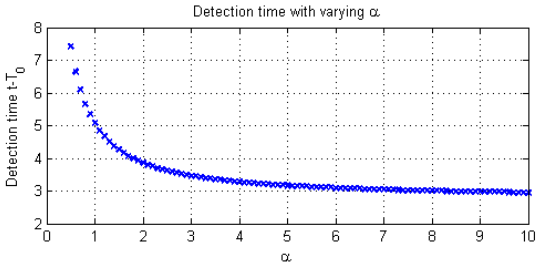


Figure 3.3: Detection time vs filter pole α.

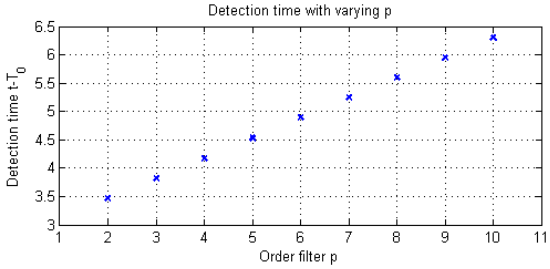


Figure 3.4: Detection time vs filter order p.

consider the case where the magnitude of the fault is $\left|\phi_I^{(k)}(x(t'),u_I(t'))\right| = M, \quad \forall\, t' \in [T_0,t]$ and $M > 2(\bar{\chi}_I^{(k)} + \bar{\epsilon}_{\xi_I}^{(k)} + \bar{\epsilon}_{\zeta_I}^{(k)})$. Then, the solution of (3.64) gives the actual detection time.

Due to the mathematical complexity, a symbolic solution of an upper bound of the detection time is obtainable only in the case of a first order filter, i.e. $p = 1$ so that the threshold is given by (3.30) or (3.31) with $H_p(s) = \frac{\alpha}{s+\alpha}$ and abrupt faults, that is $b_I \to \infty$ and therefore the term $q_2(t, T_0, \alpha) = 0$. First we consider the case where no noise or disturbances are present so that $\bar{\chi}_I^{(k)} = \bar{\eta}_I^{(k)}$ and the threshold is given by (3.30). Therefore, the upper bound of the detection time T_d can be found from solving

$$q_1(T_d, T_0, \alpha) = \frac{2(p-1)!}{M}\bar{r}_I^{(k)}(T_d),$$

which can be written as

$$\gamma\big(p, \alpha(T_d - T_0)\big) = \frac{2}{M}\bar{\chi}_I^{(k)}\gamma\big(p, \alpha T_d\big). \tag{3.73}$$

For $p = 1$ and using the equation $\gamma\big(1, \alpha t\big) = 1 - e^{-\alpha t}$, the solution of (3.73) gives the upper bound of the detection time:

$$T_d = T_0 + \frac{1}{\alpha}\ln\left(\frac{M - 2e^{-\alpha T_0}\bar{\chi}_I^{(k)}}{M - 2\bar{\chi}_I^{(k)}}\right).$$

Similarly, in the case where the threshold is given by (3.31) the upper bound of the detection time T_d can be found from solving

$$\gamma\big(p, \alpha(T_d - T_0)\big) = \frac{2(p-1)!}{M}\left(\frac{1}{(p-1)!}\big(\bar{\chi}_I^{(k)} + \alpha\xi_{I,b}^{(k)} + \zeta_{I,b}^{(k)}\big)\gamma\big(p, \alpha T_d\big) + \alpha\xi_{I,b}^{(k)}\right),$$

which for $p = 1$ has solution

$$T_d = T_0 + \frac{1}{\alpha}\ln\left(\frac{M - 2e^{-\alpha T_0}\big(\bar{\chi}_I^{(k)} + \alpha\xi_{I,b}^{(k)} + \zeta_{I,b}^{(k)}\big)}{M - 2\big(\bar{\chi}_I^{(k)} + 2\alpha\xi_{I,b}^{(k)} + \zeta_{I,b}^{(k)}\big)}\right).$$

Part (c) of the theorem shows that by increasing the value of the pole α, the upper bound on the detection time (and sometimes the actual detection time as explained before) decreases. On the other hand, the value of α regulates the cutoff frequency of the filter where

Christodoulos Kallis

the damping begins, so the pole location has an inherent trade-off between noise damping and fault detection speed. Figure 3.3 demonstrates the detection time of an incipient fault (time between fault occurrence and detection) for several values of α, obtained by solving equation (3.64) for $T_0 = 5, p = 2, b_I = 0.4, \bar{\chi}_I^{(k)} = 1, \bar{\epsilon}_{\xi_I}^{(k)} = \bar{\epsilon}_{\zeta_I}^{(k)} = 0.01, M = 3$. The figure clearly shows that while α increases the detection time decreases but after a certain value of α the detection time stays almost the same.

Part (d) of the theorem states that in the case of abrupt faults, the upper bound on the detection time increases as the order p of the filter increases. Although the proof is for the case of abrupt faults, the same behavior is observed in the case of incipient faults as well. Figure 3.4 demonstrates the detection time of an incipient fault for p ranging from 2 to 10, obtained by solving equation (3.64) for $T_0 = 5, \alpha = 3, b_I = 0.4, \bar{\chi}_I^{(k)} = 1, \bar{\epsilon}_{\xi_I}^{(k)} = \bar{\epsilon}_{\zeta_I}^{(k)} = 0.01, M = 3$ and it is clearly seen that the detection time increases as the filter order p increases. An obvious downside of higher order filtering is the possible increased detection time.

Remark 3.1. *Part (d) has also a qualitative explanation as it has necessarily to do with the phase lag introduced by the filter which increases with p. Simply put, by increasing p results in increased phase lag or delay between the input and output signal of the filter and since the detectability of a fault relies on the filtered signals, the detection time increases according to the delay incurred.*

Remark 3.2. *Prior to the occurrence of a fault, the residual differs from zero due to the effect of the filtered noise, filtered disturbances and filtered modeling uncertainty as indicated by (3.15). When a fault occurs, the residual is permanently contaminated by the filtered fault function as shown in the proof of Theorem 3.1. In general, the location of the poles simply affects the effectiveness of the noise and disturbance dampening. To make things more clear, consider Theorem 3.2 and 3.3 which rely on the special case of the filter $H_p(s)$ given in (3.37). Theorem 3.2, states that in the case of a fault (abrupt or incipient), which satisfies the conditions given in the Theorem then the fault is guaranteed to be detected. Note that this is irrespective of the location of the filters' poles. In fact, and as shown in Theorem 3.3, having faster poles results in smaller upper bound on the detection time or even smaller actual detection time. In conclusion, the location of the poles does not limit the duration of the residual activation when a fault occurs, but instead the residual is permanently affected by the filtered fault function. Therefore, the location of the poles has an inherent trade-off between noise damping and fault detection speed.*

Christodoulos Keliris

3.5 Simulation Results

In this section, we consider a numerical example based on a system of two inverted pendulums connected by a spring, as described in [78] using the modifications in [98]. The models of the two subsystems $I = 1, 2$ are given by

$$\begin{aligned}\dot{x}_I^{(1)} =& x_I^{(2)} + \zeta_I^{(1)}\\ \dot{x}_I^{(2)} =& f_I^{(2)} + g_I^{(2)} + \eta_I^{(2)} + \zeta_I^{(2)}\end{aligned}$$

where for the first subsystem $I = 1$ the nominal and interconnection functions are given by:

$$\begin{aligned}f_1^{(2)} =& \left(\frac{m_1 gr}{J_1} - \frac{kr^2}{4J_1}\right)\sin(x_1^{(1)}) + \frac{kr}{2J_1}(l-b) + \frac{u_1}{J_1}\\ g_1^{(2)} =& \frac{kr^2}{4J_1}\sin(x_2^{(2)})\end{aligned}$$

and for the second subsystem $I = 2$ the respective functions are

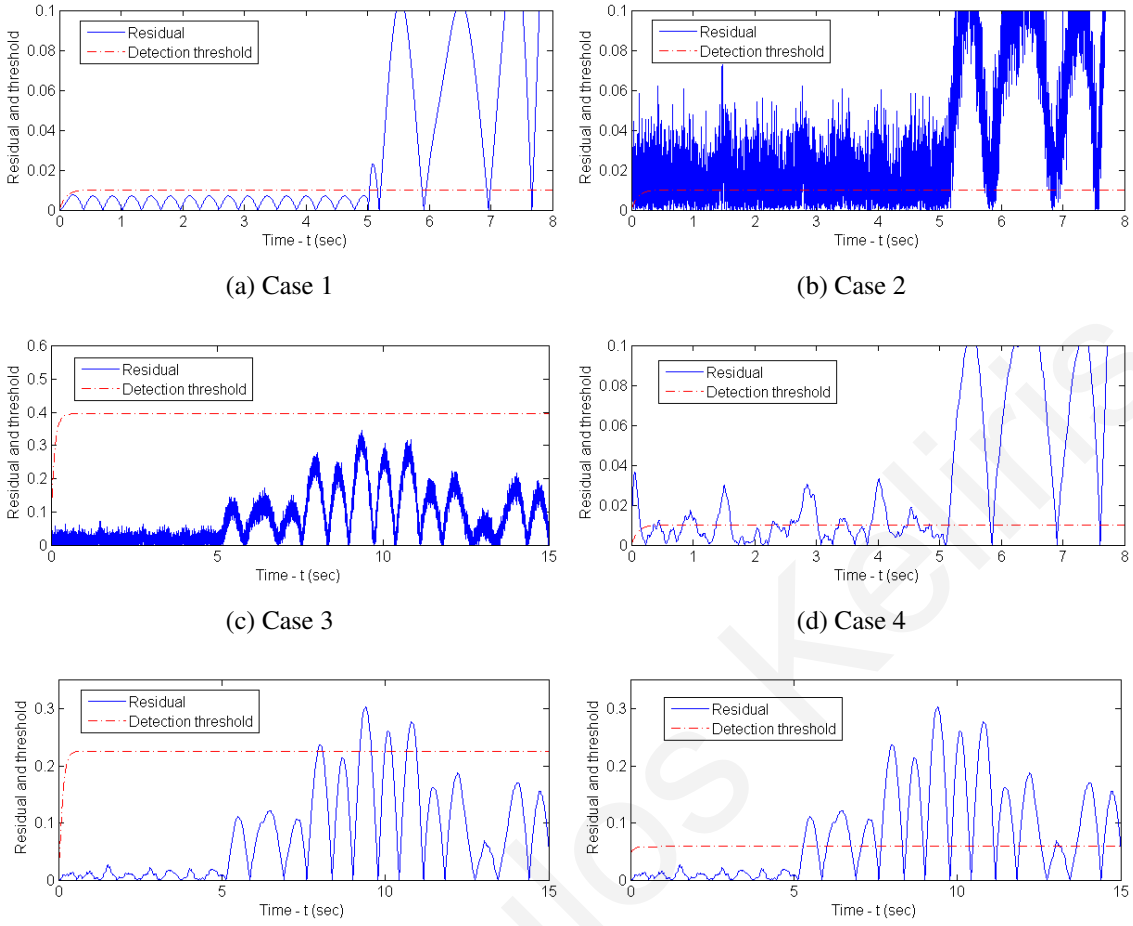
$$\begin{aligned}f_2^{(2)} =& \left(\frac{m_2 gr}{J_2} - \frac{kr^2}{4J_2}\right)\sin(x_2^{(1)}) + \frac{kr}{2J_2}(l-b) + \frac{u_2}{J_2}\\ g_2^{(2)} =& \frac{kr^2}{4J_2}\sin(x_1^{(2)}).\end{aligned}$$

The parameters that are used in the simulation are: m_1=2kg, m_2=2.5kg, J_1=0.5kg, J_2=0.625kg, k=30N/m, l=0.5m, b=0.4m and g=9.81m/s^2. The nominal functions $f_1^{(2)}$ and $f_2^{(2)}$ satisfy the Lipschitz condition with $L_{f_1^{(2)}} = 33.5$ and $L_{f_2^{(2)}} = 21$ respectively under healthy mode of operation. The interconnection terms $g_1^{(2)}$ and $g_2^{(2)}$ satisfy the Lipschitz condition with $L_{g_1^{(2)}} = 3.75$ and $L_{g_2^{(2)}} = 3$ respectively. The modeling uncertainties of both subsystems are assumed to be $\eta_1^{(2)} = \eta_2^{(2)} = 0.1\sin(10t)$, thus $\bar{\eta}_1^{(2)} = \bar{\eta}_2^{(2)} = 0.1$. The inputs u_I are derived based on a simple decentralized proportional feedback controller that stabilizes each subsystem and are given by $u_I = 20e_I$, I=1,2 where $e_I = -x_I^{(1)}$ is the tracking error. In this example, we consider an abrupt multiplicative actuator fault in subsystem 1 where the input changes to $u_1 = (1+\theta_1)\bar{u}_1$, where $\bar{u}_1$ is the nominal control input in the non-fault case and $\theta_1 \in [-1, 0]$ is the parameter characterizing the magnitude of the fault. The fault occurs at $T_0 = 5$ sec with a magnitude $\theta_1 = -0.075$.

The simulation results are demonstrated through different case scenarios which illustrate the presence of false alarms or missed detections. In all cases the residual is given by (3.14) but with different filter $H(s)$ used. The detection threshold used will be specified at the explanation of each scenario.

In "Case 1", we consider the ideal case where no disturbance or measurement noise are present so the detection threshold is given by (3.30) using a first order filter $H(s) = \frac{s}{s+10}$.

Christodoulos Keliris



(a) Case 1

(b) Case 2

(c) Case 3

(d) Case 4

(e) Case 5a

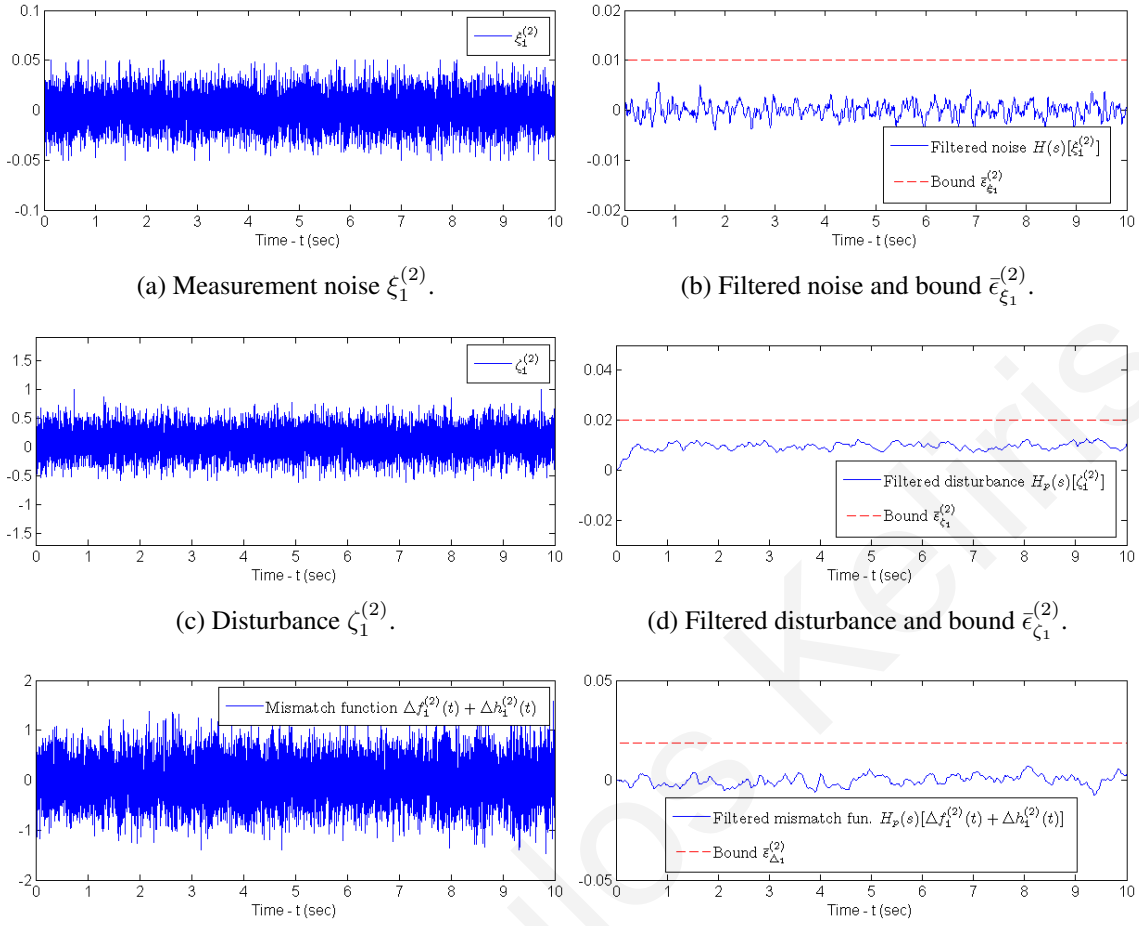
(f) Case 5b

Figure 3.5: Residual signal and fault detection threshold for Cases 1-5, under different assumptions and different selections of the filter $H(s)$.

The results of the fault detection module that monitors the measurement $y_1^{(2)}$ for faults are shown in Figure 3.5a, where it can be seen that the fault is successfully detected with no false alarms prior to the fault occurrence.

Next, the disturbance and measurement noise terms are added in order to demonstrate their effects to the detectability scheme used in "Case 1". More specifically, we consider a Gaussian-type measurement noise with mean $\mu_\xi = 0$, variance $\sigma_\xi^2 = 0.00025$ and max magnitude 0.05 (bound used $\xi_{I,b}^{(k)}$=0.05, $k = 1, 2$). The disturbance is also considered Gaussian with mean $\mu_\zeta = 0.1$, variance $\sigma_\zeta^2 = 0.05$ and max magnitude 0.9 (bound used $\zeta_{I,b}^{(k)}$=1, $k = 1, 2$). Figures 3.6a and 3.6c illustrate the measurement noise $\xi_1^{(2)}$ and disturbance $\zeta_1^{(2)}$ respectively. Therefore, in "Case 2" the measurement and noise terms are present but ignored in the fault detection scheme by using the same detection threshold (3.30) as in "Case 1". In other words, the threshold (3.30) is implemented by making use of the noisy measurements y_I instead of the actual state variables x_I and similarly the noisy $\bar{y}_I$ is used instead of $\bar{x}_I$.

Christodoulos Keliris



(a) Measurement noise $\xi_1^{(2)}$.

(b) Filtered noise and bound $\bar{\epsilon}_{\xi_1}^{(2)}$.

(c) Disturbance $\zeta_1^{(2)}$.

(d) Filtered disturbance and bound $\bar{\epsilon}_{\zeta_1}^{(2)}$.

(e) Mismatch function $\Delta f_1^{(2)}(t) + \Delta g_1^{(2)}(t)$.

(f) Filtered mismatch function and bound $\bar{\epsilon}_{\Delta_1}^{(2)}$.

Figure 3.6: Measurement noise, disturbance, mismatch function and their filtered versions.

The results are shown in Figure 3.5b where it is clearly seen that the residual significantly exceeds its detection threshold, leading to false alarms. This indicates the impact that even the slightest noise has on the fault detection performance, if it is not taken into consideration.

In "Case 3" the presence of measurement noise and disturbance are taken into account by using their respective bounds $\xi_{I,b}$ and $\zeta_{I,b}$, so the detection threshold is implemented according to (3.31), again using a first order filter $H(s) = \frac{s}{s+10}$. The results are shown in Figure 3.5c where it is seen that the threshold is never crossed, thus the fault remains undetected. Case 2 and 3 illustrate the trade-off between robustness and detectability, where the presence of measurement noise or disturbance may lead to either false alarms or missed detections, both of which are undesirable.

In "Case 4" the noisy measurements y_I are filtered with a low-pass filter $H_f(s) = \frac{50^2}{(s+50)^2}$ in order to remove noise and then the filtered measurements $y_{I,f}$ are used in the implementation of the detection threshold (3.30) with $H(s) = \frac{s}{s+10}$. In other words, the threshold (3.30) is used as in "Case 1", but now it makes use of the filtered measurements $y_{I,f}$ instead of x_I,

Christodoulos Keliris

since $y_{I,f}$ are considered noise-free. Figure 3.5d demonstrates the results of "Case 4" where it can be seen that the residual exceeds sporadically its detection threshold prior to the fault occurrence, therefore leading to false alarms.

Finally, in "Case 5" we use the proposed fault detection scheme with $H(s) = \frac{s}{s+10}\frac{50^2}{(s+50)^2}$ so that the measurement noise and disturbance are dampened. Figure 3.6b shows the filtered measurement noise $H(s)\big[\xi_1^{(2)}(t)\big]$ and the bound $\bar{\epsilon}_{\xi_1}^{(2)} = 0.01$ that is used and similarly Figure 3.6d shows the filtered disturbance $H_p(s)\big[\zeta_1^{(2)}(t)\big]$ and the bound $\bar{\epsilon}_{\zeta_1}^{(2)} = 0.02$. First, in "Case 5a" we implement the detection threshold given by (3.25) and the results are shown in Figure 3.5e, where it is shown that the fault is successfully detected at around $t = 7.5s$ with no prior false alarms. In "Case 5b" we use the less conservative detection threshold by implementing (3.29). The mismatch function $\Delta f_1^{(2)}(t) + \Delta g_1^{(2)}(t)$ is shown in Figure 3.6e and its filtered version $H_p(s)[\Delta f_1^{(2)}(t) + \Delta g_1^{(2)}(t)]$ along with the bound $\bar{\epsilon}_{\Delta_1}^{(2)}$ in Figure 3.6f. From these figures it is clearly seen that the mismatch function behaves like a random signal and therefore its filtered version is close to zero since all the rapid variations are extinguished. Finally, Figure 3.5f shows the successful fault detection in "Case 5b" which occurs at around $t = 5.5s$ which is faster in comparison to the "Case 5a", indicating the effectiveness of the proposed scheme.

3.6 Conclusion

In this chapter a distributed fault detection filtering approach for a class of interconnected, continuous-time, nonlinear systems with modeling uncertainties, measurement noise and disturbances is presented. Under certain assumptions, a distributed estimation scheme is designed and suitable adaptive detection thresholds are derived. Based on the proposed scheme, a fault detectability condition is obtained that characterizes quantitatively the class of faults that can be detected. The main contribution of this chapter is the novel design of the residual and threshold signals with built-in noise dampening filters. Intuitively, this allows dampening of the noise in a certain frequency range, thus facilitating the design of less conservative thresholds, which enhances fault detectability. The implementation is based on a distributed framework, where each subsystem is monitored by a local fault detection scheme. Each module requires the input and output measurements of the subsystem that is monitoring and also the measurements of all interconnecting subsystems that are influencing the subsystem that the specific module is monitoring. Furthermore, for a specific case of the filter $H_p(s)$, rigorous analytical results regarding the fault detectability properties, the

Christodoulos Kallis

detection time and the relation of the detection time to the filter's order and pole location were derived. This fault detectability and detection time analysis provides a useful theoretical framework for investigating the effect of the filter order and pole location on the fault detection properties.

Christodoulos Keliris

Chapter 4

A Fault Detection Filtering Approach for Input-Output Systems

In Chapter 3 the filtering approach for nonlinear fault diagnosis for the full-state measurement case was considered and, a rigorous investigation of the filtering impact (according to the poles' location and filters' order) on the detection time was presented. In this chapter, we relax the assumption of the availability of all the state measurements (through the design of a nonlinear observer) whilst maintaining the use of filters for dampening the uncertainty effects similarly to Chapter 3. In this chapter, a nonlinear observer-based approach is proposed for distributed fault detection of a class of interconnected input-output nonlinear systems, which is robust to modeling uncertainty and measurement noise. The use of nonlinear observer design allows for a particular class of input-output systems to be considered, which guarantees that the state estimation error for the nominal part of the system converges to zero. Then, the proposed fault detection scheme yields its decisions for the actual uncertain system by exploiting the designed observer characteristics. Specifically, the detection threshold construction is based on the (filtered) state estimation error for which an adaptive bound is obtained through Lyapunov analysis. In addition, a general class of filters is integrated in the design for the purpose of attenuating the measurement noise and hence it facilitates the design of tight, adaptive detection thresholds. Due to the lack of full-state measurement though, the filtering design in this chapter is treated as a linear state transformation, rather than exploiting Laplace's transform properties, which allows a more general class of filters to be considered in comparison to Chapter 3. More importantly, the fault detection approach proposed in this work makes use of nonlinear observer design which correlates the fault detection with the observer characteristics. Finally, a simplified problem formulation is

Christodoulos Kellis

investigated as a special case in which the nonlinear observer design is not needed for fault detection purposes and therefore a simplified fault detection scheme is proposed, maintaining the use of filters for dampening the measurement noise. Through the special case we derive further intuition about the key challenges that arise in the most general case. The nonlinear observer design was also not required in Chapter 3 since full state measurement was considered.

As explained in Section 3.1, the distributed fault detection scheme is based on local fault filtering schemes with each one assigned to monitor one subsystem. Each local fault detection scheme receives the input and output measurements of the subsystem it monitors and also the output measurements of all the interconnected subsystems that influence the subsystem under consideration and on the basis of this on-line information the local fault detection scheme provides a decision regarding the health of the subsystem it monitors.

The chapter is organized as follows: in Section 4.1, the problem formulation for distributed fault detection of a class of input-output nonlinear dynamical systems with modeling uncertainty and measurement noise is presented and in Section 4.2 details regarding the nonlinear observer design are given. In Section 4.3 the design of the distributed fault detection scheme based on Lyapunov analysis combined with a filtering approach is presented in detail, in Section 4.4 a fault detectability condition is derived and in Section 4.5 a special case of the nonlinear system which does not require the aforementioned nonlinear observer design is considered and a modified fault detection scheme is presented. In Section 4.6 a simulation example illustrates the concepts presented and finally, Section 4.7 provides some concluding remarks.

4.1 Problem Formulation

Consider a large-scale distributed nonlinear dynamic system, which is comprised of N subsystems Σ_I, $I \in \{1, ..., N\}$, described by:

$$\Sigma_I : \begin{cases} \dot{x}_I(t) = A_I x_I(t) + f_I\big(x_I(t), \bar{C}_I \bar{x}_I(t), u_I(t)\big) + \eta_I\big(x_I(t), \bar{x}_I(t), u_I(t), t\big) \\ \qquad\qquad +\beta_I(t - T_0)\phi_I\big(x(t), u_I(t)\big) \qquad (4.1) \\ y_I(t) = C_I x_I(t) + \xi_I(t), \qquad (4.2) \end{cases}$$

where $x_I \in \mathbb{R}^{n_I}$, $u_I \in \mathbb{R}^{m_I}$ and $y_I \in \mathbb{R}^{p_I}$ are the state, input and measured output vectors of the I-th subsystem respectively and $x \triangleq \left[x_1^\top, x_2^\top, \ldots, x_N^\top\right]^\top \in \mathbb{R}^n$ is the state vector of the overall system. The vectors $\bar{x}_I \in \mathbb{R}^{\bar{n}_I}$ and $\bar{C}_I \bar{x}_I \in \mathbb{R}^{\bar{p}_I}$ denote the state variables and the

Christodoulos Keliris

corresponding output variables, respectively, of neighboring subsystems that affect the I-th subsystem. Specifically, the interconnection variables $\bar{C}_I \bar{x}_I$ are a subset of noiseless output variables of neighboring subsystems and, this special form is required for the design of the nonlinear observer. The matrix $A_I \in \mathbb{R}^{n_I} \times \mathbb{R}^{n_I}$ and the function $f_I : \mathbb{R}^{p_I} \times \mathbb{R}^{\bar{p}_I} \times \mathbb{R}^{m_I} \mapsto \mathbb{R}^{n_I}$ characterize the known nominal function dynamics of the I-th subsystem and, are derived from the linearization at the origin of the I-th nominal nonlinear subsystem. Note that the function f_I, which, due to the linearization, contains only terms strictly higher than a linear function with respect to x_I, contains also the known part of the interconnection function between the I-th and its neighboring subsystems. The linearization at the origin for obtaining the matrix A_I and the nonlinear function f_I is required for designing the nonlinear observer, something that was first noted in [81] but, got somewhat overlooked in other research works on the design of nonlinear observers.

Moreover, note that the influence of the interconnected subsystems is known with some uncertainty (measurement noise). This is also required for the nonlinear observer design so that the measurements of the interconnection variables can be used between the fault detection agents, without requiring a more complex nonlinear observer design for guaranteeing the stability of the observers (since in this case state estimates would be required to be exchanged among the fault detection agents).

The matrix $C_I \in \mathbb{R}^{p_I} \times \mathbb{R}^{n_I}$ is the known nominal output matrix of the I-th subsystem. The vector function $\eta_I : \mathbb{R}^{n_I} \times \mathbb{R}^{\bar{n}_I} \times \mathbb{R}^{m_I} \times \mathbb{R}^{+} \mapsto \mathbb{R}^{n_I}$ denotes the modeling uncertainty associated with the nominal dynamics and $\xi_I \in \mathcal{D}_{\xi_I} \subset \mathbb{R}^{p_I}$ ($\mathcal{D}_{\xi_I}$ is a known compact set) represents the measurement noise.

Furthermore, the term $\beta_I(t - T_0)\phi_I(x, u_I)$ characterizes the time-varying fault function dynamics affecting the I-th subsystem. More specifically, the term $\phi_I : \mathbb{R}^n \times \mathbb{R}^{m_I} \mapsto \mathbb{R}^{n_I}$ represents the unknown fault function and the term $\beta_I(t - T_0) : \mathbb{R} \mapsto \mathbb{R}^{+}$ models the time evolution of the fault, where T_0 is the unknown time of the fault occurrence [70]. Note that the fault function ϕ_I may depend on the global state variable vector x and not only on the local state vector x_I. From a practical perspective, this allows for propagative fault effects to be transferred across neighboring subsystems (as it is the case in real networks such as electric power systems, transportation systems, etc). In this work, no particular modeling is considered for the time profile $\beta_I(t - T_0)$ which can be used to model both abrupt and incipient faults. Instead, we simply consider it to be zero prior to the fault occurrence, i.e. $\beta_I(t - T_0) = 0$, for all $t < T_0$. In fact, the faults can be permanent, temporary or even intermittent. As long as they are sufficiently "large" and last for sufficiently long time, then

Christodoulos Kallis

they can be detected by the proposed scheme. Moreover, the estimation of the fault function ϕ_I and of the fault occurrence time T_0 is not dealt in this work.

The objective is to design and analyze a distributed fault detection scheme, in which, for each subsystem Σ_I, a local fault detection algorithm, receiving local noisy measurements $y_I(t)$ and noisy measurements $\bar{y}_I(t)$ from neighboring subsystems is considered. It is assumed that there exist feedback controllers for selecting u_I such that some desired control objectives are achieved. The feedback controllers for generating $u_I(t)$ may be configured in a centralized framework or a distributed framework. In this work, we do not deal explicitly with the control problem, but, instead, we focus on the fault detection issue in the case of partial state measurements and in the presence of faults ϕ_I, modeling uncertainties η_I and measurement noise ξ_I. In order to deal with this task, a nonlinear observer is designed for each subsystem that guarantees that the state estimation error, for the nominal nonlinear system, converges to zero. Note that this was not required in Chapter 3 since full state measurement was assumed. Then, the designed observer is combined with filtering for attenuating measurement noise and is used in a novel way for the derivation of suitable adaptive thresholds for the filtered state estimation error of the uncertain system, which guarantee no false alarms. The novelty in the filtering approach with regards to Chapter 3 stems from its treatment as a linear state transformation rather than the exploitation of the Laplace's transform properties, which allows a more general class of filters to be considered. Therefore, the fault detection scheme is inherently tied with the nonlinear observer design and relies on its performance for the detection of faults.

The notation $\|\cdot\|$ used in this chapter denotes the Euclidean 2-norm for vectors and the matrix norm induced by the 2-norm for matrices, whereas $|\cdot|$ indicates the absolute value of a scalar function. Furthermore, $w^\top$ indicates the transpose of the vector or matrix w. The following assumptions are used throughout the chapter:

***Assumption* 4.1.** *For each subsystem Σ_I, $I \in \{1, ..., N\}$, the pair (A_I, C_I) is observable.*

***Assumption* 4.2.** *For each subsystem Σ_I, $I \in \{1, ..., N\}$, the local state variables $x_I(t)$ and the local inputs $u_I(t)$ belong to a known compact region $\mathcal{D}_{x_I}$ and $\mathcal{D}_{u_I}$ respectively before and after the occurrence of a fault, i.e. $x_I(t) \in \mathcal{D}_{x_I}$, $u_I(t) \in \mathcal{D}_{u_I}$ for all $t \geq 0$ (well-posedness).*

***Assumption* 4.3.** *The modeling uncertainty η_I in each subsystem is an unstructured and possibly unknown nonlinear function of x_I, $\bar{x}_I$, u_I and t but bounded by a known positive*

Christodoulos Kallis

function $\bar{\eta}_I$:

$$\|\eta_I(x_I, \bar{x}_I, u_I, t)\| \leq \bar{\eta}_I(y_I, \bar{y}_I, u_I), \tag{4.3}$$

for all $t \geq 0$ and for all $(x_I, \bar{x}_I, u_I) \in \mathcal{D}_I$, where $\bar{y}_I \in \mathcal{D}_{\bar{y}_I} \subset \mathbb{R}^{\bar{p}_I}$ is the measurable noisy counterpart of $\bar{C}_I\bar{x}_I$, i.e. $\bar{y}_I = \bar{C}_I\bar{x}_I + \bar{\xi}_I$, $\bar{\xi}_I \in \mathcal{D}_{\bar{\xi}_I} \subset \mathbb{R}^{\bar{p}_I}$ and $\bar{\eta}_I(y_I, \bar{y}_I, u_I) \geq 0$ is a known bounding function in some region of interest $\mathcal{D}_I = \mathcal{D}_{x_I} \times \mathcal{D}_{\bar{x}_I} \times \mathcal{D}_{u_I} \subset \mathbb{R}^{n_I} \times \mathbb{R}^{\bar{n}_I} \times \mathbb{R}^{m_I}$. The regions $\mathcal{D}_{\bar{\xi}_I}$, $\mathcal{D}_{\bar{y}_I}$ and $\mathcal{D}_I$ are known compact sets.

Assumption 4.1 is required for the design of a nonlinear observer to allow estimation of the unmeasurable state variables, while Assumption 4.2 is needed for well-posedness, since we do not consider the fault accommodation problem in this thesis. Assumption 4.3 characterizes the class of modeling uncertainties being considered. In practice, the system can be modeled more accurately in certain regions of the state space. Therefore, the fact the bound $\bar{\eta}_I$ is a function of y_I, $\bar{y}_I$ and u_I provides more flexibility by allowing the designer to take into consideration any possible prior knowledge of the system.

The nonlinear function f_I is required to satisfy the following assumption.

***Assumption* 4.4.** *The function $f_I(x_I, \bar{y}_I, u_I)$ satisfies the generalized Lipschitz condition:*

$$\|f_I(z_1, \bar{y}_I, u_I) - f_I(z_2, \bar{y}_I, u_I)\| \leq \|G_{f_I}(z_1 - z_2)\|, \tag{4.4}$$

for $z_1, z_2 \in \mathcal{D}_{x_I}, \bar{y}_I \in \mathcal{D}_{\bar{y}_I}, u_I \in \mathcal{D}_{u_I}$ where $G_{f_I} \in \mathbb{R}^{n_I \times n_I}$ is a known constant matrix.

The generalized condition stated in Assumption 4.4 is less conservative than the standard Lipschitz condition:

$$\|f_I(z_1, \bar{y}_I, u_I) - f_I(z_2, \bar{y}_I, u_I)\| \leq \gamma_{f_I}\|z_1 - z_2\|, \tag{4.5}$$

where γ_{f_I} is the Lipschitz constant for the function $f_I(x_I, \bar{y}_I, u_I)$ with respect to x_I in the region $\mathcal{D}_{x_I}$. Although the subsequent analysis is based on (4.4), the standard Lipschitz condition (4.5) can be used by replacing G_{f_I} with $\gamma_{f_I}I$ in what follows. The main difference between the generalized condition in Assumption 4.4 and the standard Lipschitz condition is that the matrix G_{f_I} can be sparse (for example owing to some a-priori information on f_I) and therefore $\|G_{f_I}(z_1 - z_2)\|$ can be significantly smaller than $\gamma_{f_I}\|z_1 - z_2\|$.

As discussed in Section 3.1 and illustrated in Figure 3.1, in general, the distributed fault detection scheme is composed of N local filtered fault detection agents $\mathcal{F}_I$, one for each subsystem Σ_I. Each local fault detection agent $\mathcal{F}_I$ requires the input and output measurements of the subsystem Σ_I that is monitoring and also the measurements of all interconnecting

Christodoulos Keliris

subsystems Σ_J that are influencing Σ_I. Note that these last measurements are communicated by neighboring fault detection agents $\mathcal{F}_J$, and not by the subsystems Σ_J. Therefore, there is the need of communication between the fault detection modules depending on their interconnections. Note that, the interconnection variables $\bar{C}_I\bar{x}_I$ are measurable with some uncertainty as $\bar{y}_I = \bar{C}_I\bar{x}_I + \bar{\xi}_I$. Hence, the fault detection modules exchange these measurements and are used by the nonlinear observer for the state estimation and for the generation of the residual and threshold signals. Therefore, the distributed nature of the scheme stems from the fact that there is communication between the fault detection modules depending on their interconnections. The example shown in Figure 4.1 illustrates the distributed fault detection scheme for the case of three subsystems Σ_1, Σ_2, Σ_3, where Σ_1 affects Σ_2 and Σ_2 affects Σ_3.

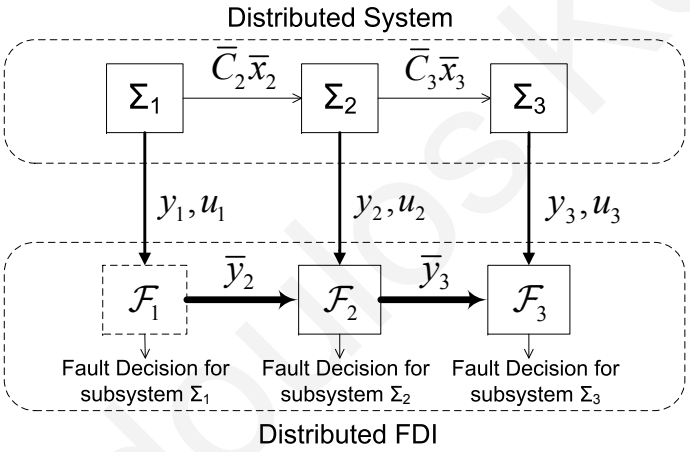


Figure 4.1: Distributed fault detection scheme for the case of three subsystems Σ_1, Σ_2, Σ_3, where Σ_1 affects Σ_2 and Σ_2 affects Σ_3.

To dampen the effects of measurement uncertainties $\xi_I(t)$, each measured variable $y_I^{(k)}(t)$ (k-th component of y_I) is filtered by $H_p(s)$, where $H_p(s)$ is a p-th order, strictly proper, asymptotically stable transfer function

$$H_p(s) = \frac{d_{p-1}s^{p-1} + d_{p-2}s^{p-2} + \ldots + d_0}{s^p + c_{p-1}s^{p-1} + \ldots + c_1 s + c_0}. \tag{4.6}$$

Note that, the filter used in filtering the output measurements in this case (given by (4.6)) is more general than the one used in the full state measurement case in the previous Chapter which was given by (3.5). The expansion in the class of filters that can be used in this Chapter, stems from the treatment of the filtering as a linear state transformation. Note that, a different filter $H_p(s)$ can be used by each fault detection agent $\mathcal{F}_I$, but in this work, without loss of generality, we consider the same filter for simplicity. Within $\mathcal{F}_I$, the same filter must be used though.

Christodoulos Keliris

The local fault detection module $\mathcal{F}_I$ is made of the local nonlinear observer, the local residual signal and the corresponding detection threshold. The local nonlinear observer is given by

$$\dot{\hat{x}}_I(t) = A_I\hat{x}_I(t) + f_I\big(\hat{x}_I(t), \bar{y}_I(t), u_I(t)\big) + L_I\big(y_I(t) - \hat{y}_I(t)\big) \tag{4.7}$$

$$\hat{y}_I(t) = C_I\hat{x}_I(t), \tag{4.8}$$

where $L_I \in \mathbb{R}^{n_I \times p_I}$ is the observer gain matrix (to be designed) with $\hat{x}_I(0) = 0$ for simplicity. The local residual vector signal $r_I(t)$ is generated according to

$$r_I(t) \triangleq H_p(s)\big[y_I(t) - \hat{y}_I(t)\big]. \tag{4.9}$$

The decision regarding the detection of a fault is made when $\|r_I(t)\| \geq \bar{r}_I(t)$ in any local fault detection agent $\mathcal{F}_I$, where $\bar{r}_I(t)$ is the detection threshold (to be specified later on). To summarize, for each subsystem Σ_I the proposed fault filtered scheme $\mathcal{F}_I$ relies on designing a nonlinear observer gain matrix L_I (Section 4.2), selecting a filter with transfer function $H_p(s)$ of the form of (4.6) and obtaining a computable adaptive detection threshold $\bar{r}_I(t)$ (Section 4.3). Figure 4.2 illustrates the implementation of the local filtered fault detection scheme $\mathcal{F}_I$ resulting from equations (4.6)-(4.9).

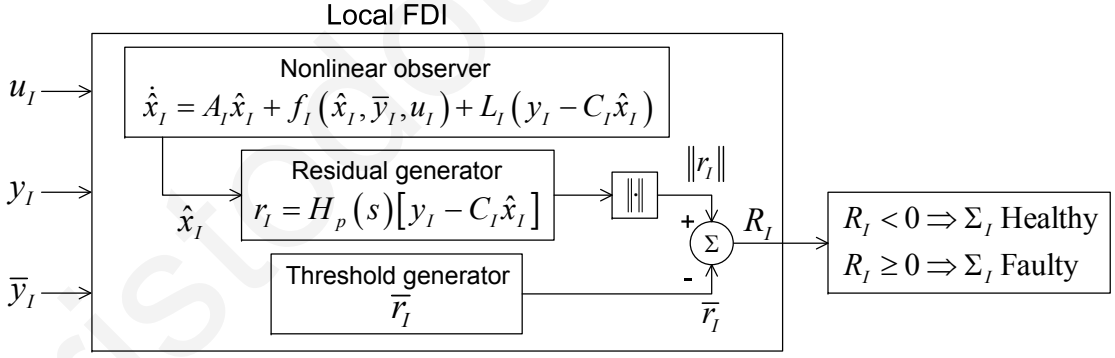


Figure 4.2: Local Filtered Fault Detection Scheme $\mathcal{F}_I$.

The filter $H_p(s)$ given by (4.6) is asymptotically stable and hence BIBO stable. Therefore, for bounded measurement noise $\xi_I(t)$, the filtered measurement noise $\epsilon_{\xi_I}(t) \triangleq H_p(s)\,[\xi_I(t)]$ is bounded:

$$\|\epsilon_{\xi_I}(t)\| \leq \bar{\epsilon}_{\xi_I}(t), \quad \forall t \geq 0,$$

where $\bar{\epsilon}_{\xi_I}(t)$ is assumed to be a known bounding function. The approach presented in this chapter relies on the knowledge of the bounds for the filtered measurement noise and the modeling uncertainty, as in Chapter 3, in order to guarantee no false alarms.

Christodoulos Keliris

4.2 Nonlinear Observer Design

In this section, we deal with the problem of the design of the local nonlinear observer so that the state estimation error for the nominal part of the nonlinear subsystem Σ_I under healthy mode of operation converges to zero exponentially fast. Then, the fault detection scheme will make use of filtering and rely on the performance of the nonlinear observer in order to establish an adaptive bound on the filtered state estimation error for the uncertain and with measurement noise system Σ_I and hence, obtain a suitable detection threshold.

The nonlinear observer design problem, although well investigated, still presents some open challenges. Let us consider the simple case in which the nonlinear system is described by

$$\dot{x} = Ax + f(x, u)$$
$$y = Cx$$

where (A, C) is observable and $f(x, u)$ is a nonlinearity that contains only high-order (larger than 1) terms with respect to x and, is considered locally Lipschitz with respect to x in a compact region of interest $x \in \mathcal{D}_x$ and has a Lipschitz constant γ, i.e. $\|f(x, u) - f(\hat{x}, u)\| \leq \gamma\|x - \hat{x}\|$. The observer is considered of the form

$$\dot{\hat{x}} = A\hat{x} + f(\hat{x}, u) + L(y - C\hat{x})$$

and therefore the estimation error $\tilde{x} \triangleq x - \hat{x}$ satisfies

$$\dot{\tilde{x}} = (A - LC)\tilde{x} + f(x, u) - f(\hat{x}, u).$$

In this case, the seminal paper by Thau [81] gives a necessary condition for observer stability without providing a design procedure to select L. Simply selecting L for $(A - LC)$ to be Hurwitz (eigenvalues in the open left half complex plane) is not sufficient to guarantee that the estimation error will converge to zero. In fact, necessary and sufficient conditions for the asymptotic stability of the observers' estimation error dynamics are given by Rajamani [72]. Therefore, the *fault detection application* for this class of systems is more difficult.

In this work, the proposed fault detection scheme relies on the performance of the nonlinear observer design in order to expand the class of input-output systems that can be considered. More specifically, the nonlinear observer design is achieved by modifying a design method proposed by Phanomchoeng & Rajamani [68]. More information regarding the modification will be given later on.

Christodoulos Kellis

First, we rewrite the dynamics of Σ_I given in (4.1) as:

$$\begin{aligned}\dot{x}_I(t) =& A_I x_I(t) + f_I\big(x_I(t), \bar{y}_I(t), u_I(t)\big) + \Delta f_I(t) + \eta_I\big(x_I(t), \bar{x}_I(t), u_I(t), t\big) \\ &+ \beta_I(t - T_0)\phi_I\big(x(t), u_I(t)\big)\end{aligned} \tag{4.10}$$

where $\Delta f_I(t) \triangleq f_I\big(x_I(t), \bar{C}_I\bar{x}_I(t), u_I(t)\big) - f_I\big(x_I(t), \bar{y}_I(t), u_I(t)\big)$.

It is worth noting that in (4.10), the functions Δf_I, η_I and ϕ_I are in general unknown. By excluding these functions, we define the known nominal system as follows:

$$\dot{x}_{I,N}(t) = A_I x_{I,N}(t) + f_I\big(x_{I,N}(t), \bar{y}_I(t), u_I(t)\big) \tag{4.11}$$

$$y_{I,N}(t) = C_I x_{I,N}(t). \tag{4.12}$$

Essentially, we first design a nonlinear observer so that the nominal state estimation error $\tilde{x}_{I,N}$ converges to zero exponentially fast by satisfying the condition $\dot{V}_I(\tilde{x}_{I,N}) < -\mu_I V_I(\tilde{x}_{I,N})$ for some $\mu_I > 0$. Obviously, the state estimation error $\tilde{x}_I$ for the actual nonlinear uncertain system given by (4.10) does not converge to zero as in the case of the nominal system. But, by using the designed observer and Lyapunov theory we are able to derive adaptive bounds for the state estimation error $\tilde{x}_I$ in the fault-free case. In addition, filtering is used in order to mitigate the effects of noise and derive tighter bounds for the filtered state estimation error $\tilde{x}_{I,f}$.

The nonlinear observer given in (4.7), (4.8) is constructed based on the dynamics described by (4.11), (4.12). By using (4.11) and (4.7), the nominal state estimation error $\tilde{x}_{I,N} \triangleq x_{I,N} - \hat{x}_I$ (based on the nominal system so that $y_I = y_{I,N}$ in (4.7)) satisfies the nominal error dynamics

$$\dot{\tilde{x}}_{I,N} = A_{I,0}\tilde{x}_{I,N} + f_I(x_{I,N}, \bar{y}_I, u_I) - f_I(\hat{x}_I, \bar{y}_I, u_I), \tag{4.13}$$

where $A_{I,0} \triangleq A_I - L_I C_I$.

By using the Lyapunov function candidate $V_I(\tilde{x}_{I,N}) = \tilde{x}_{I,N}^\top P_I \tilde{x}_{I,N}$ where $P_I > 0$ and $P_I \in \mathbb{R}^{n_I \times n_I}$, its time derivative along the trajectories of the nominal error dynamics (4.13) satisfies:

$$\dot{V}_I(\tilde{x}_{I,N}) = \tilde{x}_{I,N}^\top \big[A_{I,0}^\top P_I + P_I A_{I,0}\big]\tilde{x}_{I,N} + 2\tilde{x}_{I,N}^\top P_I \big[f_I(x_{I,N}, \bar{y}_I, u_I) - f_I(\hat{x}_I, \bar{y}_I, u_I)\big].$$

In matrix form, the previous equation can be written as $\dot{V}_I(\tilde{x}_{I,N}) = w_I^\top Q_I w_I$ where

$$w_I^\top \triangleq \begin{bmatrix}\tilde{x}_{I,N}^\top & (f_I(x_{I,N}, \bar{y}_I, u_I) - f_I(\hat{x}_I, \bar{y}_I, u_I))^\top\end{bmatrix},$$

$$Q_I^\top \triangleq \begin{bmatrix}A_{I,0}^\top P_I + P_I A_{I,0} & P_I \\ P_I & 0\end{bmatrix}.$$

Christodoulos Keliris

The proposed fault detection scheme (to be presented in Section 4.3) relies on designing a nonlinear observer such that the nominal error dynamics satisfy $\dot{V}_I(\tilde{x}_{I,N}) < -\mu_I V_I(\tilde{x}_{I,N})$ for some given $\mu_I > 0$. It must be pointed out though, that any nonlinear observer of the form (4.7), (4.8) satisfying the aforementioned condition can be used for the fault detection task. Below, we propose such an observer by modifying the one described in [68] by adapting it to our problem formulation and by adding one extra degree of freedom (μ_I), which is required for the fault detection task.

Lemma 4.1. *Consider the nominal nonlinear system described in (4.11), (4.12) satisfying the generalized Lipschitz assumption (4.4) and the nonlinear observer described in (4.7), (4.8). The nominal state estimation error $\tilde{x}_{I,N}$ dynamics are asymptotically stable if and only if an observer gain matrix L_I can be chosen such that for some positive definite symmetric matrix P_I, $\mu_I \geq 0$, $\epsilon_I \geq 0$ the following LMI is satisfied*

$$\begin{bmatrix} A_{I,0}^\top P_I + P_I A_{I,0} + \mu_I P_I + \epsilon_I G_{f_I}^\top G_{f_I} & P_I \\ P_I & -\epsilon_I I \end{bmatrix} < 0 \tag{4.14}$$

where $A_{I,0} = A_I - L_I C_I$.

Proof. Suppose there exist matrices L_I and P_I which satisfy (4.14). Let this choice of L_I be used in the observer (4.7) for state estimation of the nominal system given by (4.11). According to the S-Procedure Lemma [68], let $V_0(x)$ and $V_1(x)$ be two arbitrary quadratic forms over $\mathbb{R}^n$. Then $V_0(x) < 0$ for all $x \in \mathbb{R}^n - \{0\}$ satisfying $V_1(x) \leq 0$ if and only if there exist $\epsilon \geq 0$ such that $V_0(x) < \epsilon V_1(x)$ for all $x \in \mathbb{R}^n - \{0\}$. Applying the S-Procedure Lemma with $V_0(x) = \dot{V}_I(\tilde{x}_{I,N})$ and $V_1(x) = -V_I(\tilde{x}_{I,N})$ we find that $\dot{V}_I(\tilde{x}_{I,N}) < 0$ if and only if there exists $\mu_I \geq 0$ such that $\dot{V}_I(\tilde{x}_{I,N}) < -\mu_I V_I(\tilde{x}_{I,N})$. Next, by rewriting $V_I(\tilde{x}_{I,N})$ in matrix form as

$$V_I(\tilde{x}_{I,N}) = w_I^\top \begin{bmatrix} P_I & 0 \\ 0 & 0 \end{bmatrix} w_I$$

then $W_{I,0}(\tilde{x}_{I,N}) \triangleq \dot{V}_I(\tilde{x}_{I,N}) + \mu_I V_I(\tilde{x}_{I,N})$, can be written in matrix form as

$$W_{I,0}(\tilde{x}_{I,N}) = w_I^\top R_I w_I \tag{4.15}$$

where

$$R_I \triangleq \begin{bmatrix} A_{I,0}^\top P_I + P_I A_{I,0} + \mu_I P_I & P_I \\ P_I & 0 \end{bmatrix}.$$

Christodoulos Keliris

Now, consider the nominal nonlinear system (4.11) which satisfies the generalized Lipschitz condition (4.4). The generalized Lipschitz condition (4.4) can be rewritten as

$$W_{I,1}(\tilde{x}_{I,N}) \triangleq -\|(G_{f_I}\tilde{x}_{I,N})\|^2 + \|f_I(x_{I,N}, \bar{y}_I, u_I) - f_I(\hat{x}_I, \bar{y}_I, u_I)\|^2 \leq 0,$$

or in matrix form as

$$W_{I,1}(\tilde{x}_{I,N}) = w_I^\top \begin{bmatrix} -G_{f_I}^\top G_{f_I} & 0 \\ 0 & I \end{bmatrix} w_I \leq 0. \tag{4.16}$$

Applying the S-Procedure Lemma to (4.15) and (4.16) we find that $W_{I,0}(\tilde{x}_{I,N}) < 0$ if and only if there exists $\epsilon_I \geq 0$ such that $W_{I,0}(\tilde{x}_{I,N}) < \epsilon_I W_{I,1}(\tilde{x}_{I,N})$. Hence, by rewriting the last inequality in matrix form using (4.15) and (4.16), the necessary and sufficient condition for observer design is given by $\exists \mu_I \geq 0, \epsilon_I \geq 0$ such that (4.14) holds. □

As stated in [68], (4.14) is a LMI involving two unknown matrices L_I and P_I. The LMI can be replaced by a Riccati inequality in just one variable P_I. A necessary and sufficient condition so that the observers' estimation error converges to zero can be obtained in terms of the existence of a solution to a certain Riccati inequality and is formalized in the following Lemma.

Lemma 4.2. *Consider the nominal nonlinear system described in (4.11), (4.12) which satisfies the generalized Lipschitz condition (4.4). Then, there exists a nonlinear observer as described in (4.7), (4.8) such that the state estimation error $\tilde{x}_{I,N}$ is quadratically stabilized if and only if there exist $\mu_I \geq 0$, $\epsilon_I > 0$ and $\delta_I \in \mathbb{R}$ such that the following Riccati inequality has a symmetric, positive definite solution P_I:*

$$\bar{A}_I^\top P_I + P_I \bar{A}_I + \epsilon_I G_{f_I}^\top G_{f_I} + \frac{P_I P_I}{\epsilon_I} - \delta_I^2 C_I^\top C_I < 0 \tag{4.17}$$

where $\bar{A}_I \triangleq A_I + \frac{1}{2}\mu_I I$. The observer gain matrix can then be chosen as $L_I = \frac{\delta_I^2}{2} P_I^{-1} C_I^\top$.

Proof. The proof follows along the lines of [68] but it is included here for completeness. By applying the Schur inequality [68] to (4.14), the necessary and sufficient condition for observer design is

$$A_{I,0}^\top P_I + P_I A_{I,0} + \mu_I P_I + \epsilon_I G_{f_I}^\top G_{f_I} + \frac{1}{\epsilon_I} P_I P_I < 0$$

which can be rewritten as

$$\bar{A}_I^\top P_I + P_I \bar{A}_I + \epsilon_I G_{f_I}^\top G_{f_I} + \frac{1}{\epsilon_I} P_I P_I - C_I^\top L_I^\top P_I - P_I L_I C_I < 0. \tag{4.18}$$

Christodoulos Keliris

Then by using $L_I = \frac{\delta_I^2}{2} P_I^{-1} C_I^\top$, (4.18) is written as (4.17) and hence (4.17) is a necessary condition. The sufficiency is proved as follows. Suppose there exists a positive definite solution P_I to (4.17) and let the observer gain matrix be chosen as $L_I = \frac{\delta_I^2}{2} P_I^{-1} C_I^\top$. Then, by replacing $C_I^\top$ in (4.17) with $C_I^\top = \frac{2}{\delta_I^2} P_I L_I$, (4.17) can be rewritten as (4.18) and therefore (4.14) is satisfied. □

Lemma 4.2 provides a simple way to obtain an appropriate observer gain matrix L_I through the solution of a Riccati inequality. According to this design, the nominal error satisfies $\dot{V}_I(\tilde{x}_{I,N}) < -\mu_I V_I(\tilde{x}_{I,N})$ for some given $\mu_I \geq 0$ and hence it converges to zero. For fault detection purposes it is required that $\mu_I > 0$; after some mathematical manipulations the nominal state estimation error satisfies:

$$\|\tilde{x}_{I,N}(t)\| < \alpha_I e^{-\frac{\mu_I}{2} t} \|x_{I,N}(0)\| \tag{4.19}$$

where $\alpha_I \triangleq \sqrt{\frac{\lambda_{max}(P_I)}{\lambda_{min}(P_I)}}$. Note that $\tilde{x}_{I,N}(0) = x_{I,N}(0)$ because $\hat{x}_I(0)$ is considered zero for simplicity.

4.3 Distributed Fault Detection

In this section, we exploit the nonlinear observer properties derived in the previous section for the task of fault detection. The observer design is based on the known nominal system dynamics, not taking into account the unknown components; i.e., the uncertainty term η_I, the function discrepancy $\Delta f_I(t)$ and the measurement noise ξ_I, which affect the performance of the observer. In particular, the measurement noise term may create significant challenges in fault diagnosis since it is difficult to strike a balance between conservative detection thresholds and minimizing the presence of false alarms.

In the following, $z_f(t) \triangleq H_p(s)\big[z(t)\big]$ denotes the filtered version of any signal $z(t)$ which is passed through a filter with transfer function $H_p(s)$.

Let $h_p(t)$ be the impulse response associated with $H_p(s)$; i.e. $h_p(t) \triangleq \mathcal{L}^{-1}\left[H_p(s)\right]$. Then $z_f(t)$ can be written as $z_f(t) = \int_0^t h_p(\tau) z(t-\tau)\, \mathrm{d}\tau$. By taking the derivative of $z_f(t)$ and using the Leibniz integral rule, we obtain

$$\dot{z}_f(t) = H_p(s)\left[\dot{z}(t)\right] + h_p(t) z(0). \tag{4.20}$$

By using (4.13) and (4.20) with $z = \tilde{x}_{I,N}$, the filtered nominal estimation error $\tilde{x}_{I,N,f}$

Christodoulos Kellis

satisfies:

$$\begin{aligned}\dot{\tilde{x}}_{I,N,f}(t) &= H_p(s)\big[\dot{\tilde{x}}_{I,N}(t)\big] + h_p(t)\tilde{x}_{I,N}(0)\\ &= H_p(s)\big[A_{I,0}\tilde{x}_{I,N}(t) + f_I\big(x_{I,N}(t),\bar{y}_I(t),u_I(t)\big) - f_I\big(\hat{x}_I(t),\bar{y}_I(t),u_I(t)\big)\big]\\ &\quad + h_p(t)x_{I,N}(0)\\ &= A_{I,0}\tilde{x}_{I,N,f}(t) + H_p(s)\big[f_I\big(x_{I,N}(t),\bar{y}_I(t),u_I(t)\big) - f_I\big(\hat{x}_I(t),\bar{y}_I(t),u_I(t)\big)\big]\\ &\quad + h_p(t)x_{I,N}(0). \end{aligned} \tag{4.21}$$

It is worth noting that the term $x_{I,N}(0)h_p(t)$ in (4.21) converges to zero because of Lemma 3.3.

In the following analysis we will prove that the filtered nominal state estimation error $\tilde{x}_{I,N,f}$ converges to zero with exponential speed. This will be used as the basis for the derivation of an adaptive detection threshold for the filtered state estimation error $\tilde{x}_{I,f}(t) = H_p(s)[\tilde{x}_I(t)]$, where $\tilde{x}_I \triangleq x_I - \hat{x}_I$. By using Lemma 3.3 and (4.19), the norm of $\tilde{x}_{I,N,f}(t)$ satisfies:

$$\begin{aligned}\|\tilde{x}_{I,N,f}(t)\| &= \|H_p(s)\big[\tilde{x}_{I,N}(t)\big]\|\\ &\leq \int_0^t |h_p(\tau)|\|\tilde{x}_{I,N}(t-\tau)\|\,\mathrm{d}\tau \qquad (4.22)\\ &< \int_0^t \kappa e^{-\upsilon\tau}\alpha_I e^{-\frac{\mu_I}{2}(t-\tau)}\|x_{I,N}(0)\|\,\mathrm{d}\tau. \qquad (4.23)\end{aligned}$$

By solving the integral we obtain:

$$\|\tilde{x}_{I,N,f}(t)\| < \begin{cases}\alpha_I\kappa\|x_{I,N}(0)\|\frac{e^{-\upsilon t}-e^{-\frac{\mu_I}{2}t}}{\frac{\mu_I}{2}-\upsilon} & \text{if } \upsilon \neq \frac{\mu_I}{2},\\ \alpha_I\kappa\|x_{I,N}(0)\|te^{-\frac{\mu_I}{2}t} & \text{if } \upsilon = \frac{\mu_I}{2}.\end{cases}$$

Note that $\lim_{t\to\infty}\|\tilde{x}_{I,N,f}(t)\| = 0$ for any $\mu_I, \upsilon > 0$ and in fact the filtered nominal state estimation error $\tilde{x}_{I,N,f}$ converges to zero exponentially fast; i.e. there exists a sufficiently small $\rho_I > 0$ such that

$$\|\tilde{x}_{I,N,f}(t)\| < \alpha_I e^{-\frac{\rho_I}{2}t}\|x_{I,N}(0)\|. \tag{4.24}$$

In other words, since the nominal state estimation error dynamics $\tilde{x}_{I,N}$ converge to zero exponentially fast with a convergence speed of $\frac{\mu_I}{2}$ (see (4.19)), the filtered nominal state estimation error $\tilde{x}_{I,N,f}$ still converges to zero exponentially fast and we need to find a sufficient convergence rate $\frac{\rho_I}{2}$, something that is dealt with later. In addition, note that since the filtered nominal state estimation error $\tilde{x}_{I,N,f}$ decays exponentially, it satisfies a differential inequality of the form

$$\dot{V}_I(\tilde{x}_{I,N,f}) < -\rho_I V_I(\tilde{x}_{I,N,f}) \tag{4.25}$$

Christodoulos Keliris

with initial condition $\tilde{x}_{I,N,f}(0) = x_{I,N}(0)$. This is similar to the case of the nominal state estimation error $\tilde{x}_{I,N}$ which satisfies $\dot{V}_I(\tilde{x}_{I,N}) < -\mu_I V_I(\tilde{x}_{I,N})$ and (4.19).

Therefore, by combining (4.22), (4.19) and (4.24), a sufficient condition for (4.25) to be satisfied for all $t > t^*$ is:

$$\int_0^t |h_p(\tau)| e^{-\frac{\mu_I}{2}(t-\tau)} \, \mathrm{d}\tau < e^{-\frac{\rho_I}{2}t} \quad \forall t > t^*. \tag{4.26}$$

Thus, ρ_I must be selected so that (4.26) is satisfied. The time $t^* \geq 0$ is selected by the designer and constitutes a time instant for which the detection results in the time interval $[0, t^*]$ are ignored and its purpose is to allow greater flexibility in choosing ρ_I. As it is clearly indicated by (4.26), the selection of a suitable value for ρ_I depends on the filters' impulse response $h_p(t)$ and the nonlinear observer design (due to μ_I).

In order to select ρ_I, we first select a time $t^* \geq 0$. Then, for a given filter (hence $h_p(t)$) and a given nonlinear observer design that satisfies the required condition $\dot{V}_I(\tilde{x}_{I,N}) < -\mu_I V_I(\tilde{x}_{I,N})$ (hence given μ_I), we verify for various values of ρ_I that (4.26) holds for all $t > t^*$ through numerical methods by trying to solve (4.26) as equality and verifying that no solution exists; i.e. the two sides of (4.26) do not cross for $t > t^*$. Actually, the left part of (4.26) can be bounded by an exponentially decaying function due to the asymptotic stability of the filter (i.e. see Lemma 3.3) and hence, there always exists a sufficiently small exponential decay speed ρ_I. Having (4.26) satisfied for all $t > t^*$ rather than $t > 0$ provides greater flexibility in choosing ρ_I since the left and right side of (4.26) are allowed to cross at time t^* and hence (4.25) is satisfied for all $t > t^*$. This comes at the mere cost of having to initiate the fault detection scheme at time zero but ignoring the detection results in the time interval $[0, t^*]$ since false alarms may occur during this time (because (4.25) might not hold for $t \leq t^*$). In any case, no false alarms are guaranteed and the only downside of using t^* is the possibility of delayed fault detection at a time after t^* in the case that a fault occurs before t^*.

Note that, an easier to check but more conservative condition than (4.26) (in the sense that the values of ρ_I satisfying it are more limited) is obtained by using (4.26) and Lemma 3.3 and results in the following condition to be satisfied for all $t > t^*$:

$$\begin{aligned} \kappa \frac{e^{-\upsilon t} - e^{-\frac{\mu_I}{2}t}}{\frac{\mu_I}{2} - \upsilon} &\leq e^{-\frac{\rho_I}{2}t} \quad \text{if } \upsilon \neq \frac{\mu_I}{2} \\ \kappa t e^{-\frac{\mu_I}{2}t} &\leq e^{-\frac{\rho_I}{2}t} \quad \text{if } \upsilon = \frac{\mu_I}{2}. \end{aligned} \tag{4.27}$$

Note that, by using $V_I(\tilde{x}_{I,N,f}) = \tilde{x}_{I,N,f}^\top P_I \tilde{x}_{I,N,f}$, its time derivative along the trajectories

Christodoulos Keliris

of the filtered nominal error dynamics (4.21) satisfies:

$$\begin{aligned}\dot{V}_I(\tilde{x}_{I,N,f}) = \tilde{x}_{I,N,f}^\top \big[A_{I,0}^\top P_I + P_I A_{I,0}\big]\tilde{x}_{I,N,f} \\ + 2\tilde{x}_{I,N,f}^\top P_I \Big[H_p(s)\big[f_I\big(x_{I,N}(t), \bar{y}_I(t), u_I(t)\big) \\ - f_I\big(\hat{x}_I(t), \bar{y}_I(t), u_I(t)\big)\big] + h_p(t) x_{I,N}(0)\Big]\end{aligned}$$

and, by using (4.25), we have

$$\begin{aligned}2\tilde{x}_{I,N,f}^\top P_I \Big[H_p(s)\big[f_I\big(x_{I,N}(t), \bar{y}_I(t), u_I(t)\big) - f_I\big(\hat{x}_I(t), \bar{y}_I(t), u_I(t)\big)\big] + h_p(t) x_{I,N}(0)\Big] \\ + \tilde{x}_{I,N,f}^\top \big[A_{I,0}^\top P_I + P_I A_{I,0}\big]\tilde{x}_{I,N,f} < -\rho_I V_I(\tilde{x}_{I,N,f}).\end{aligned} \tag{4.28}$$

Now, let's consider the uncertain nonlinear system (4.10), (4.2) with the designed nonlinear observer given in (4.7), (4.8). In the absence of a fault, the state estimation error $\tilde{x}_I$ satisfies:

$$\dot{\tilde{x}}_I(t) = A_{I,0}\tilde{x}_I(t) + \tilde{\Delta} f_I(t) + \Delta f_I(t) + \eta_I\big(x_I(t), \bar{x}_I(t), u_I(t), t\big) - L_I \xi_I(t), \tag{4.29}$$

where $\tilde{\Delta} f_I(t) \triangleq f_I\big(x_I(t), \bar{y}_I(t), u_I(t)\big) - f_I\big(\hat{x}_I(t), \bar{y}_I(t), u_I(t)\big)$. By using (4.29) and (4.20) with $z = \tilde{x}_I$, the filtered state estimation error $\tilde{x}_{I,f}$ satisfies:

$$\begin{aligned}\dot{\tilde{x}}_{I,f}(t) &= H_p(s)\big[\dot{\tilde{x}}_I(t)\big] + h_p(t)\tilde{x}_I(0) \\ &= H_p(s)\big[A_{I,0}\tilde{x}_I(t) + \tilde{\Delta} f_I(t) + \Delta f_I(t) - L_I\xi_I(t) + \eta_I\big(x_I(t), \bar{x}_I(t), u_I(t), t\big)\big] \\ &\quad + h_p(t) x_I(0) \\ &= A_{I,0}\tilde{x}_{I,f}(t) + \tilde{\Delta} f_{I,f}(t) + h_p(t) x_I(0) + \chi_I(t)\end{aligned} \tag{4.30}$$

where

$$\chi_I(t) \triangleq H_p(s)\left[\eta_I(x_I(t), \bar{x}_I(t), u_I(t), t)\right] + \epsilon_{\Delta_I}(t) - L_I \epsilon_{\xi_I}(t) \tag{4.31}$$

$$\epsilon_{\Delta_I}(t) \triangleq H_p(s)\left[\Delta f_I(t)\right]. \tag{4.32}$$

Therefore, the time derivative of $V_I(\tilde{x}_{I,f})$ along the trajectories of the filtered error dynamics (4.30) satisfies:

$$\dot{V}_I(\tilde{x}_{I,f}) = 2\tilde{x}_{I,f}^\top P_I\big(\tilde{\Delta} f_{I,f}(t) + h_p(t) x_I(0)\big) + \tilde{x}_{I,f}^\top \big[A_{I,0}^\top P_I + P_I A_{I,0}\big]\tilde{x}_{I,f} + 2\tilde{x}_{I,f}^\top P_I \chi_I(t). \tag{4.33}$$

Note that, the first two terms in (4.33) are as in (4.28) and hence, (4.33) can be re-written as

$$\dot{V}_I(\tilde{x}_{I,f}) < -\rho_I V_I(\tilde{x}_{I,f}) + 2\tilde{x}_{I,f}^\top P_I \chi_I(t) \tag{4.34}$$

Christodoulos Keliris

with initial condition $\tilde{x}_{I,f}(0) = x_I(0)$.

For the derivation of the detection threshold, we will use the following assumption.

***Assumption* 4.5.** *The filtered function mismatch term $\epsilon_{\Delta_I}(t)$ is bounded by a computable positive function $\bar{\epsilon}_{\Delta_I}(t)$; i.e.,*

$$\|\epsilon_{\Delta_I}(t)\| \leq \bar{\epsilon}_{\Delta_I}(t), \quad \forall t \geq 0. \tag{4.35}$$

Assumption 4.5 is based on the observation that filtering dampens the error effect of measurement noise present in the function mismatch term $\Delta f_I(t)$. A suitable selection of the bound $\bar{\epsilon}_{\Delta_I}$ can be made through the use of simulations (i.e. Monte Carlo methods) by filtering the function mismatch term $\Delta f_I(t)$ using the known nominal function dynamics and the available noise characteristics (recall that the measurement noise is assumed to take values in a compact set).

Now, consider the term $\chi_I(t)$ which satisfies:

$$\begin{aligned}\|\chi_I(t)\| &= \|H_p(s)\left[\eta_I(x_I, \bar{y}_I, u_I, t)\right] + \epsilon_{\Delta_I}(t) - L_I\epsilon_{\xi_I}(t)\| \\ &\leq \int_0^t |h_p(t-\tau)|\|\eta_I(x_I(\tau), \bar{y}_I(\tau), u_I(\tau), \tau)\|\,\mathrm{d}\tau + \bar{\epsilon}_{\Delta_I}(t) + \|L_I\|\bar{\epsilon}_{\xi_I}(t). \end{aligned} \tag{4.36}$$

Based on (4.36), the designed bounding function $\bar{\chi}_I(t)$ such that $\|\chi_I(t)\| \leq \bar{\chi}_I(t)$ is given by:

$$\bar{\chi}_I(t) \triangleq \bar{H}_p(s)\left[\bar{\eta}_I(y_I(t), \bar{y}_I(t), u_I(t))\right] + \|L_I\|\bar{\epsilon}_{\xi_I}(t) + \bar{\epsilon}_{\Delta_I}(t), \tag{4.37}$$

where $\bar{H}_p(s)$ is a transfer function with impulse response $\bar{h}_p(t)$ that satisfies $|h_p(t)| \leq \bar{h}_p(t)$ for all $t \geq 0$ and can be determined by using a method of Section 3.2.2.

In the following, we proceed to "decouple" the last term of (4.34). For this, we will use the following lemma.

Lemma 4.3. *[88]. For any $z_1, z_2 \in \mathbb{R}^n$ and any positive definite matrix $P \in \mathbb{R}^{n\times n}$ the following inequality holds:*

$$2z_I^\top z_2 \leq z_1^\top P z_1 + z_2^\top P^{-1} z_2.$$

Using Lemma 4.3, for any positive scalar γ_I we can write the sign-indefinite term $2\tilde{x}_{I,f}^\top P_I \chi_I(t)$ in (4.34) using $z_1 = \sqrt{\gamma_I}\tilde{x}_{I,f}$, $z_2 = \frac{1}{\sqrt{\gamma_I}} P_I\chi_I(t)$ and $P = P_I$ as follows:

$$\begin{aligned} 2\tilde{x}_{I,f}^\top P_I\chi_I(t) \leq& \gamma_I \tilde{x}_{I,f}^\top P_I \tilde{x}_{I,f} + \gamma_I^{-1}\chi_I^\top(t) P_I \chi_I(t) \\ \leq& \gamma_I V_I(\tilde{x}_{I,f}) + \gamma_I^{-1}\lambda_{max}(P_I)\bar{\chi}_I^2(t). \end{aligned} \tag{4.38}$$

Christodoulos Keliris

Therefore, (4.34) can be written as

$$\dot{V}_I(\tilde{x}_{I,f}) + (\rho_I - \gamma_I)V_I(\tilde{x}_{I,f}) < \gamma_I^{-1}\lambda_{max}(P_I)\bar{\chi}_I^2(t). \tag{4.39}$$

Next, we utilize the following Lemma [43].

Lemma 4.4. *[43]. Let $w(t), V(t) : [0, \infty) \mapsto \mathbb{R}$. Then, for any finite α, if $\dot{V}(t) \leq -\alpha V(t) + w(t)$ for all $t \geq t_0 \geq 0$ then it implies that*

$$V(t) \leq e^{-\alpha(t-t_0)}V(t_0) + \int_{t_0}^{t} e^{-\alpha(t-\tau)}w(\tau)\,\mathrm{d}\tau \quad \forall t \geq t_0 \geq 0.$$

Therefore, by using Lemma 4.4, (4.39) becomes

$$V_I(\tilde{x}_{I,f}) < e^{-(\rho_I-\gamma_I)t}V_I(\tilde{x}_{I,f}(0)) + \int_0^t e^{-(\rho_I-\gamma_I)(t-\tau)}\gamma_I^{-1}\lambda_{max}(P_I)\bar{\chi}_I^2(\tau)\,\mathrm{d}\tau.$$

By utilizing linear filtering techniques, the previous inequality can be written as

$$V_I(\tilde{x}_{I,f}) < e^{-(\rho_I-\gamma_I)t}V_I(\tilde{x}_{I,f}(0)) + \frac{1}{s + (\rho_I - \gamma_I)}\big[\gamma_I^{-1}\lambda_{max}(P_I)\bar{\chi}_I^2(t)\big].$$

Hence, by using $\lambda_{min}(P_I)\|\tilde{x}_{I,f}\|^2 \leq V_I(\tilde{x}_{I,f}) \leq \lambda_{max}(P_I)\|\tilde{x}_{I,f}\|^2$ we obtain an adaptive bound $\tilde{x}_{I,f}^{max}(t)$ such that $\|\tilde{x}_{I,f}(t)\| < \tilde{x}_{I,f}^{max}(t)$ and is given by

$$\tilde{x}_{I,f}^{max}(t) \triangleq \alpha_I\Big(e^{-(\rho_I-\gamma_I)t}\bar{x}_{I,d}^2 + \frac{\gamma_I^{-1}}{s + (\rho_I - \gamma_I)}\big[\bar{\chi}_I^2(t)\big]\Big)^{\frac{1}{2}}, \tag{4.40}$$

where $\bar{x}_{I,d}$ is a bounding estimate of $x_I(0)$ such that $\|x_I(0)\| \leq \bar{x}_{I,d}$ and γ_I is selected so that $0 < \gamma_I < \rho_I$ so that the filter is stable.

Using (4.2) and (4.8), the residual (4.9) satisfies

$$\begin{aligned}\|r_I(t)\| &= \|H_p(s)\big[C_I\tilde{x}_I(t) + \xi_I(t)\|\big] \\ &\leq \|C_I\|\|\tilde{x}_{I,f}(t)\| + \|\epsilon_{\xi_I}(t)\|.\end{aligned}$$

Finally, by using the bound on $\|\tilde{x}_{I,f}(t)\|$, the detection threshold $\bar{r}_I(t)$ so that $\|r_I(t)\| < \bar{r}_I(t)$ is given by:

$$\bar{r}_I(t) \triangleq \|C_I\|\tilde{x}_{I,f}^{max}(t) + \bar{\epsilon}_{\xi_I}. \tag{4.41}$$

The following theorem summarizes the findings of the above analysis.

Theorem 4.1. *Consider the nominal nonlinear system described in (4.11), (4.12) satisfying the generalized Lipschitz assumption (4.4) and the nonlinear observer described in (4.7), (4.8) that is designed such that the nominal state estimation error $\tilde{x}_{I,N}$ is quadratically stabilized satisfying $\dot{V}_I(\tilde{x}_{I,N}) < -\mu_I V_I(\tilde{x}_{I,N})$ for some $\mu_I > 0$, where $V_I(\tilde{x}_{I,N}) = \tilde{x}_{I,N}^\top P_I \tilde{x}_{I,N}$,*

Christodoulos Keliris

$P_I > 0$ and $P_I \in \mathbb{R}^{n_I \times n_I}$. Then, for the uncertain nonlinear system given in (4.1), (4.2) the residual $r_I(t)$ in (4.9) which is implemented using a filter $H_p(s)$ of the form given in (4.6), satisfies $\|r_I(t)\| < \bar{r}_I(t) \quad \forall t \in [t^*, T_0)$, where $\bar{r}_I(t)$ is the detection threshold given by (4.41) and (4.40), $\bar{\chi}_I(t)$ is given by (4.37), ρ_I and t^* are selected so that (4.26) is satisfied and γ_I (in (4.40)) is a scalar selected such that $0 < \gamma_I < \rho_I$.

Remark 4.1. *The proposed approach is comprised of two main elements: the nonlinear observer design for each subsystem and the derivation of suitable threshold signals. The problem of nonlinear observer design can be considered on an individual basis for each subsystem, which means that the observer designs are essentially decoupled among them. These designs are made once, prior to the initiation of the fault detection scheme and remain the same afterwards. On the other hand, the residual and threshold generation is made in real time and the proposed fault detection scheme is distributed (but not decentralized) since there is the need of communication between the fault detection modules depending on their interconnections.*

Remark 4.2. *In the particular nonlinear observer approach (see Lemma 4.2) the constants ϵ_I, δ_I, μ_I appear. Note that, the parameters ϵ_I, δ_I affect only the nonlinear observer design and can be chosen arbitrarily as long as (4.17) can be solved whereas the choice of μ_I is more important because it affects both the nonlinear observer design and the fault detection scheme. Qualitatively, we would like μ_I to have a sufficiently large value to allow for more flexibility in choosing ρ_I later through (4.26). The fault detection scheme also depends on the choice of the parameters ρ_I and γ_I. Qualitatively, it is preferable that ρ_I is chosen sufficiently large so that γ_I can be chosen more freely according to $0 < \gamma_I < \rho_I$ in order to guarantee filter stability. Although the choice of μ_I is fixed according to the nonlinear observer design, one can find a range of values for the parameters ρ_I and γ_I so that these two can be changed online during the fault detection task.*

Remark 4.3. *In the proposed scheme, the filters are used primarily to mitigate the effects of noise and their use allows less conservative detection thresholds to be obtained. For instance, let's examine the derivation of the residual and threshold signals without the use of filtering. The residual in this case is simply $r_I(t) = y_I(t) - \hat{y}_I(t)$. By designing a nonlinear observer as described in Section 4.2, we have that the state estimation error $\tilde{x}_I$ satisfies (derived similarly to (4.34)) $\dot{V}_I(\tilde{x}_I) < -\mu_I V_I(\tilde{x}_I) + 2\tilde{x}_I^\top P_I v_I(t)$, with initial condition $\tilde{x}_I(0) = x_I(0)$ and $v_I(t) \triangleq \eta_I(x_I, \bar{x}_I, u_I, t) + \Delta f_I(t) - L_I \xi_I(t)$. In this case, a suitable*

Christodoulos Keliris

bounding function $\bar{v}_I(t)$ such that $\|v_I(t)\| \leq \bar{v}_I(t)$ is given by:

$$\bar{v}_I(t) \triangleq \bar{\eta}_I(y_I, \bar{y}_I, u_I, t) + \overline{\Delta f}_I + \|L_I\|\xi_{I,b} \tag{4.42}$$

where $\|\xi_I(t)\| \leq \xi_{I,b}$ and

$$\overline{\Delta f}_I \triangleq \sup_{\substack{(x_I, \bar{x}_I, u_I) \in \mathcal{D}_I \\ \bar{\xi}_I \in \mathcal{D}_{\bar{\xi}_I}}} \|f_I\big(x_I(t), \bar{C}_I \bar{x}_I(t), u_I(t)\big) - f_I\big(x_I(t), \bar{C}_I \bar{x}_I(t) + \bar{\xi}_I(t), u_I(t)\big)\|.$$

Since the regions $\mathcal{D}_I$ and $\mathcal{D}_{\bar{\xi}_I}$ are compact, the supremum above is finite. In addition, note that the bound $\bar{v}_I(t)$ in (4.42) depends on t because of the bounding function $\bar{\eta}_I$. In this case a suitable detection threshold (derived similarly to (4.41)) is given by:

$$\bar{r}_I(t) \triangleq \|C_I\|\tilde{x}_I^{max}(t) + \xi_{I,b} \tag{4.43}$$

where

$$\tilde{x}_I^{max}(t) \triangleq \alpha_I \Big(e^{-(\mu_I - \theta_I)t} \bar{x}_{I,d}^2 + \frac{\theta_I^{-1}}{s + (\mu_I - \theta_I)} \big[\bar{v}_I^2(t)\big] \Big)^{\frac{1}{2}}$$

is the bound on $\|\tilde{x}_I(t)\|$ prior to the fault occurrence; i.e., $\|\tilde{x}_I(t)\| < \tilde{x}_I^{max}(t)$ and the scalar θ_I is selected so that $0 < \theta_I < \mu_I$. Note that, the detection threshold in (4.43) contains terms regarding the bounds of the measurement noise ξ_I (which is multiplied by $\|L_I\|$) and of the function discrepancy term Δf_I. Therefore, the detection threshold is more conservative, which may lead to missed faults (false negatives). This is further illustrated in the simulation results.

Remark 4.4. *Note that, a different bound on the filtered state estimation error $\tilde{x}_{I,f}$ (instead of* (4.40)*) is given by $\bar{H}_p(s)\left[\tilde{x}_I^{max}(t)\right]$ (where $\tilde{x}_I^{max}(t)$ is given in Remark 4.3). However, this bound does not exploit the filtering benefits for dampening the measurement noise and its effects in order to obtain tight detection thresholds. The treatment of the filtering in this work as a linear state transformation counteracts this problem, but at the same time creates some additional challenges, that are successfully tackled, such as to how to properly select ρ_I and "correlate" the exponential convergence to zero of $\tilde{x}_{I,N,f}$ with the boundedness of $\tilde{x}_{I,f}$ (see* (4.34)*).*

4.4 Fault detectability

So far, the design and analysis was based on devising suitable thresholds $\bar{r}_I(t)$ such that in the absence of any fault we have $\|r_I(t)\| < \bar{r}_I(t)$. In the following, a fault detectability condition of the aforementioned fault detection scheme is presented, which provides a quantitative characterization of a class of faults detectable by the proposed scheme.

Christodoulos Keliris

Theorem 4.2. *Consider the nonlinear system (4.1), (4.2) with the distributed fault detection scheme described in (4.6), (4.7), (4.8), (4.9) and the detection threshold (4.41). A fault in the I-th subsystem initiated at $t = T_0$ is detectable if the following inequality is satisfied for some $\theta_I \in (0, \mu_I)$ at some time $t > T_0$:*

$$\begin{aligned}\| \int_{T_0}^{t} C_I e^{A_{I,0}(t-\tau)} \phi_{I,f}\big(x(\tau), u_I(\tau), \tau\big)\, \mathrm{d}\tau \| > \bar{r}_I(t) + \bar{\epsilon}_{\xi_I}(t) \\ + \int_0^t \|C_I e^{A_{I,0}(t-\tau)}\| \bigg(\|G_{f_I}\| \bar{H}_p(s)[\tilde{x}_I^{max,\phi}(\tau)] + |h_p(\tau)| \bar{x}_{I,d} + \bar{\chi}_I(\tau) \bigg)\, \mathrm{d}\tau\end{aligned}$$

where $\phi_{I,f}\big(x(t), u_I(t), t\big) \triangleq H_p(s)[\beta_I(t - T_0)\phi_I(x(t), u_I(t))]$ *and*

$$\tilde{x}_I^{max,\phi}(t) \triangleq \alpha_I \bigg(e^{-(\mu_I - \theta_I)t} \bar{x}_{I,d}^2 + \frac{\theta_I^{-1}}{s + (\mu_I - \theta_I)} \big[(\bar{v}_I(t) + \|\phi_{I,f}\big(x(t), u_I(t), t\big)\|)^2 \big] \bigg)^{\frac{1}{2}}.$$

Proof. At first, let us consider the state estimation error without filtering. After the occurrence of the fault, $\tilde{x}_I$ satisfies

$$\dot{V}_I(\tilde{x}_I) < -\mu_I V_I(\tilde{x}_I) + 2\tilde{x}_I^\top P_I(v_I(t) + \beta_I(t - T_0)\phi_I(x(t), u_I(t))).$$

Following the same steps as in Remark 4.3, we obtain that after the occurrence of the fault, $\tilde{x}_I$ satisfies $\|\tilde{x}_I(t)\| < \tilde{x}_I^{max,\phi}(t)$ where $\tilde{x}_I^{max,\phi}(t)$ is given in Theorem 4.2. Now, consider the filtered state estimation error $\tilde{x}_{I,f}(t)$. After the occurrence of the fault, (4.30) becomes:

$$\dot{\tilde{x}}_{I,f}(t) = A_{I,0}\tilde{x}_{I,f}(t) + \tilde{\Delta} f_{I,f}(t) + h_p(t)x_I(0) + \chi_I(t) + \phi_{I,f}\big(x(t), u_I(t), t\big). \tag{4.44}$$

The solution of (4.44), is given by:

$$\begin{aligned}\tilde{x}_{I,f}(t) = \int_0^t e^{A_{I,0}(t-\tau)} \big(\tilde{\Delta} f_{I,f}(\tau) + h_p(\tau)x_I(0) + \chi_I(\tau) \big)\, \mathrm{d}\tau \\ + \int_0^t e^{A_{I,0}(t-\tau)} \phi_{I,f}\big(x(\tau), u_I(\tau), \tau\big)\, \mathrm{d}\tau\end{aligned} \tag{4.45}$$

where initial condition term is omitted since $\tilde{x}_{I,f}(0) = 0$ due to the filters' zero initial condition. By using $r_I(t) = C_I x_{I,f}(t) + \epsilon_{\xi_I}(t)$, (4.45) and the triangle inequality, $r_I(t)$ satisfies:

$$\begin{aligned}\|r_I(t)\| \geq - \int_0^t \|C_I e^{A_{I,0}(t-\tau)}\| \|\tilde{\Delta} f_{I,f}(\tau)\|\, \mathrm{d}\tau - \|\epsilon_{\xi_I}(t)\| \\ - \int_0^t \|C_I e^{A_{I,0}(t-\tau)}\| \big(\|h_p(\tau)x_I(0)\| + \|\chi_I(\tau)\| \big)\, \mathrm{d}\tau \\ + \| \int_0^t C_I e^{A_{I,0}(t-\tau)} \phi_{I,f}\big(x(\tau), u_I(\tau), \tau\big)\, \mathrm{d}\tau \|.\end{aligned} \tag{4.46}$$

Using the generalized Lipschitz condition (4.4) and after some mathematical manipulations, the mismatch function term in the first integral of (4.46) satisfies:

$$\begin{aligned}\|\tilde{\Delta} f_{I,f}(\tau)\| &= \|H_p(s)[f_I(x_I(t), \bar{y}_I(t), u_I(t)) - f_I(\hat{x}_I(t), \bar{y}_I(t), u_I(t))]\| \\ &\leq \bar{H}_p(s)[\|G_{f_I}\tilde{x}_I(t)\|] \\ &\leq \|G_{f_I}\| \bar{H}_p(s)[\tilde{x}_I^{max,\phi}(t)],\end{aligned}$$

Christodoulos Keliris

and therefore (4.46) becomes:

$$\|r_I(t)\| \geq -\int_0^t \|C_I e^{A_{I,0}(t-\tau)}\| \Big(\|G_{f_I}\| \bar{H}_p(s)[\tilde{x}_I^{max,\phi}(t)] + |h_p(\tau)|\bar{x}_{I,d} + \bar{\chi}_I(\tau) \Big) \, \mathrm{d}\tau - \bar{\epsilon}_{\xi_I}(t)$$
$$+ \| \int_{T_0}^t C_I e^{A_{I,0}(t-\tau)} \phi_{I,f}\big(x(\tau), u_I(\tau), \tau\big) \, \mathrm{d}\tau \|.$$

For fault detection, the inequality $\|r_I(t)\| > \bar{r}_I(t)$ must hold at some time $t > T_0$, so the final fault detectability condition given in Theorem 4.2 is obtained. □

The above fault detectability theorem implicitly characterizes the type of faults that can be detected by the proposed distributed fault detection scheme. Clearly, the fault functions $\phi_I(x, u_I)$ are typically unknown and therefore this condition cannot be checked apriori. However, it provides useful intuition about the type of faults that are detectable. The use of filtering is of crucial importance in order to derive tight detection thresholds that guarantee no false alarms. As it can be seen in the detectability condition given in Theorem 4.2, the detection of the fault depends on the filtered fault function $\phi_{I,f}$. As a result, the selection of the filter plays a crucial role to the proposed scheme. Because in practice the fault function is usually comprised of "lower frequency" components, the fault function is not affected that much by filtering in comparison to the noise, thus allowing the derivation of tighter detection thresholds. A rigorous investigation of the filtering impact (according to the poles' location and filters' order) on the detection time was presented in Chapter 3.

4.5 A special case

In this section we consider a special case of the nonlinear system (4.1), (4.2), where $f_I(\cdot)$ is a function of $C_I x_I$ that can be measured with uncertainty as y_I. Earlier, $f_I(\cdot)$ was a function of x_I which could only be partially available (and again measured with uncertainty). Therefore, in this special case all the variables $C_I x_I$, $\bar{C}_I \bar{x}_I$ of the nonlinearity $f_I(\cdot)$ are measurable with some uncertainty as y_I, $\bar{y}_I$ respectively and, a key consequence is that there is no need to follow the design procedure for calculating L_I as in Section 4.2. More specifically the system dynamics are given by

$$\Sigma_I : \begin{cases} \dot{x}_I(t) = A_I x_I(t) + f_I(C_I x_I(t), \bar{C}_I \bar{x}_I(t), u_I(t)) + \eta_I(x_I(t), \bar{x}_I(t), u_I(t), t) \\ \qquad\qquad +\beta_I(t - T_0)\phi_I(x(t), u_I(t)) \qquad (4.47) \\ y_I(t) = C_I x_I(t) + \xi_I(t). \qquad (4.48) \end{cases}$$

The Assumptions that are used for this case are Assumptions 1-3 stated in Section 4.1. It must be noted though, that in this Section the notation $\|\cdot\|$ indicates any norm and not just

Christodoulos Kallis

the 2-norm used in Sections 4.1-4.4.

In this special case, we design an alternative distributed fault detection filtering scheme and obtain some additional analytical results. For simplicity, the initial conditions of the estimation model are set to zero; i.e., $\hat{x}_I(0) = 0$. Moreover, the simplified system dynamics given by (4.47), (4.48) allow for simpler mathematical calculations so that particular emphasis is given on the benefits of the proposed fault detection filtering scheme which are justified by the simulation results. Therefore, through the special case we derive further intuition about the key challenges that arise in the most general case.

In this special case, for each subsystem Σ_I, we design a local estimation model:

$$\dot{\hat{x}}_I(t) = A_I\hat{x}_I(t) + f_I(y_I(t), \bar{y}_I(t), u_I(t)) + L_I(y_I(t) - \hat{y}_I(t)) \tag{4.49}$$

$$\hat{y}_I(t) = C_I\hat{x}_I(t), \tag{4.50}$$

where the gain matrix L_I is computed so that $(A_I - L_IC_I)$ is Hurwitz (following Assumption 4.1).

The residual signal in this case is given as before by (4.9) and the detection decision of a fault in the overall system is made when $|r_I^{(k)}(t)| > \bar{r}_{I,2}^{(k)}(t)$ for at least one component $k = 1, 2, \ldots, p_I$, in any local subsystem Σ_I, where $\bar{r}_{I,2}^{(k)}(t)$ is the detection threshold for this special case (to be designed later on). Therefore, for each local subsystem and for each measured variable $y_I^{(k)}$, $k = 1, \ldots, p_I$, a residual $r_I^{(k)}$ and its corresponding threshold $\bar{r}_I^{(k)}$ is generated. Note that in the previous case in Section 4.3 there was only a single residual and threshold signal for each local subsystem (see, for instance Figure 4.2). This is because the detection threshold in the general case relies on a bound for the norm of the filtered state estimation error; i.e. $\tilde{x}_{I,f}^{max}(t)$. Hence, it is more appropriate to take into consideration the combined effect of all the available measurements for the computation of the residual (as $\|r_I(t)\|$) rather than having a residual signal (as $|r_I^{(k)}(t)|$) for each measurement $y_I^{(k)}$.

Let us consider the time interval $[0, T_0)$, prior to the occurrence of any fault. In this time interval, using (4.47)-(4.50), the state estimation error $\tilde{x}_I(t) \triangleq x_I(t) - \hat{x}_I(t)$ satisfies:

$$\begin{aligned}\dot{\tilde{x}}_I = A_{I,0}\tilde{x}_I + f_I\big(C_Ix_I(t), \bar{C}_I\bar{x}_I(t), u_I(t)\big) - f_I\big(y_I(t), \bar{y}_I(t), u_I(t)\big) \\ + \eta_I\big(x_I(t), \bar{x}_I(t), u_I(t), t\big) - L_I\xi_I(t).\end{aligned} \tag{4.51}$$

Using (4.51) and (4.20) with $z = \tilde{x}_I$, the filtered state estimation error $\tilde{x}_{I,f}$ satisfies:

$$\dot{\tilde{x}}_{I,f}(t) = H_p(s)\big[\dot{\tilde{x}}_I(t)\big] + h_p(t)\tilde{x}_I(0). \tag{4.52}$$

Christodoulos Keliris

Moreover, by using (4.51) and $\tilde{x}_I(0) = x_I(0)$ (since $\hat{x}_I(0) = 0$), (4.52) becomes:

$$\begin{aligned}\dot{\tilde{x}}_{I,f}(t) &= A_{I,0}\tilde{x}_{I,f} + H_p(s)\big[f_I\big(C_I x_I(t), \bar{C}_I \bar{x}_I(t), u_I(t)\big) - f_I\big(y_I(t), \bar{y}_I(t), u_I(t)\big)\big] \\ &\quad + H_p(s)\big[\eta_I\big(x_I(t), \bar{x}_I(t), u_I(t), t\big)\big] - L_I\epsilon_{\xi_I}(t) + h_p(t)x_I(0) \\ &= A_{I,0}\tilde{x}_{I,f}(t) + \chi_{I,2}(t), \end{aligned} \tag{4.53}$$

where

$$\begin{aligned}\chi_{I,2}(t) &\triangleq H_p(s)\left[\eta_I(x_I, \bar{x}_I, u_I, t)\right] - L_I\epsilon_{\xi_I}(t) + \epsilon_{\Delta_{I,2}}(t) + h_p(t)x_I(0), \\ \epsilon_{\Delta_{I,2}}(t) &\triangleq H_p(s)\left[\Delta f_{I,2}(t)\right], \\ \Delta f_{I,2}(t) &\triangleq f_I\big(C_I x_I(t), \bar{C}_I \bar{x}_I(t), u_I(t)\big) - f_I\big(y_I(t), \bar{y}_I(t), u_I(t)\big).\end{aligned}$$

The solution of (4.53) is

$$\tilde{x}_{I,f}(t) = e^{A_{I,0}t}\tilde{x}_{I,f}(0) + \int_0^t e^{A_{I,0}(t-\tau)}\chi_{I,2}(\tau)\,\mathrm{d}\tau. \tag{4.54}$$

***Assumption* 4.6.** *The filtered function mismatch term $\epsilon_{\Delta_{I,2}}(t)$ is bounded by a computable positive function $\bar{\epsilon}_{\Delta_{I,2}}(t)$; i.e., for all $t \geq 0$,*

$$\|\epsilon_{\Delta_{I,2}}(t)\| \leq \bar{\epsilon}_{\Delta_{I,2}}(t).$$

Similarly to Assumption 4.5 in Section 4.3, Assumption 4.6 is based on the observation that filtering dampens the error effect of measurement noise present in the function mismatch term $\Delta f_{I,2}(t)$.

Using (4.48) and (4.50) the residual (4.9) satisfies $r_I(t) = C_I\tilde{x}_{I,f}(t) + \epsilon_{\xi_I}(t)$. Using (4.54) and noting that $\tilde{x}_{I,f}(0) = 0$ due to the filters' zero initial condition, the residual becomes

$$r_I(t) = \int_0^t C_I e^{A_{I,0}(t-\tau)}\chi_{I,2}(\tau)\,\mathrm{d}\tau + \epsilon_{\xi_I}(t). \tag{4.55}$$

Using the triangle inequality, the k-th element of $r_I(t)$, i.e. $r_I^{(k)}(t)$ satisfies

$$\begin{aligned}|r_I^{(k)}(t)| &\leq \left|\int_0^t C_I^{(k)} e^{A_{I,0}(t-\tau)}\chi_I(\tau)\,\mathrm{d}\tau\right| + |\epsilon_{\xi_I}^{(k)}(t)| \\ &\leq \int_0^t \|C_I^{(k)} e^{A_{I,0}(t-\tau)}\|\|\chi_{I,2}(\tau)\|\,\mathrm{d}\tau + |\epsilon_{\xi_I}^{(k)}(t)| \end{aligned} \tag{4.56}$$

where $C_I^{(k)}$ indicates the k-th row of the matrix C_I. Note that since $A_{I,0}$ is Hurwitz there exist positive constants $\nu_I^{(k)}$, $\lambda_I^{(k)}$ such that $\|C_I^{(k)} e^{A_{I,0}t}\| \leq \nu_I^{(k)} e^{-\lambda_I^{(k)}t}$. In addition, using the fact that $|\epsilon_{\xi_I}^{(k)}(t)| \leq \|\epsilon_{\xi_I}(t)\| \leq \bar{\epsilon}_{\xi_I}(t)$, (4.56) becomes

$$|r_I^{(k)}(t)| \leq \int_0^t \nu_I^{(k)} e^{-\lambda_I^{(k)}(t-\tau)}\|\chi_{I,2}(\tau)\|\,\mathrm{d}\tau + \bar{\epsilon}_{\xi_I}(t). \tag{4.57}$$

Christodoulos Keliris

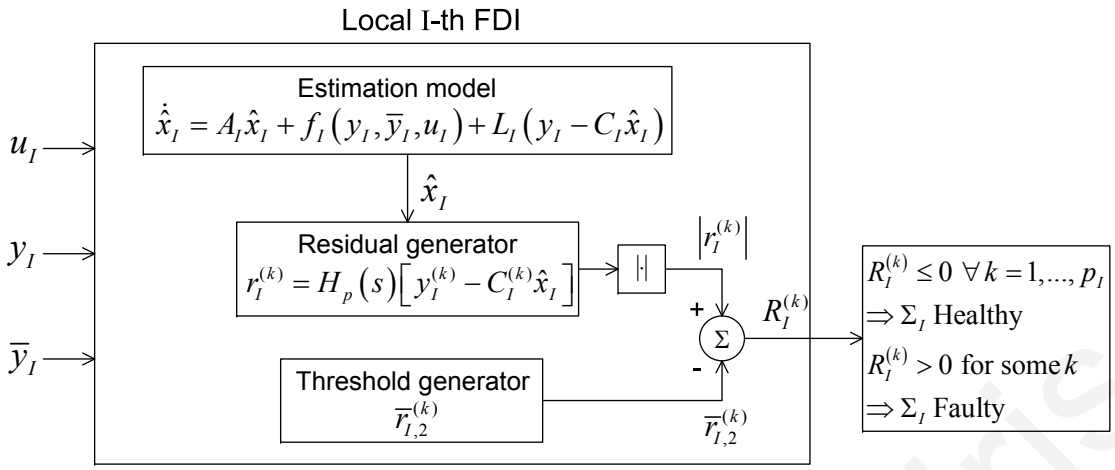


Figure 4.3: Local Filtered Fault Detection Scheme for special case.

Now, consider the term $\chi_{I,2}(t)$ which satisfies

$$\begin{aligned}\|\chi_{I,2}(t)\| &= \|H_p(s)\left[\eta_I(x_I, \bar{x}_I, u_I, t)\right] - L_I\epsilon_{\xi_I}(t) + \epsilon_{\Delta_{I,2}}(t) + h_p(t)x_I(0)\| \\ &\le \int_0^t |h_p(t-\tau)|\|\eta_I(x_I(\tau), \bar{x}_I(\tau), u_I(\tau), \tau)\|\,\mathrm{d}\tau + \|L_I\epsilon_{\xi_I}(t)\| + \|\epsilon_{\Delta_{I,2}}(t)\| \\ &\quad + |h_p(t)|\|x_I(0)\|.\end{aligned}$$

Therefore, a bounding function $\bar{\chi}_{I,2}(t)$ such that $\|\chi_{I,2}(t)\| < \bar{\chi}_{I,2}(t)$ is designed as

$$\bar{\chi}_{I,2}(t) \triangleq \quad \bar{H}_p(s)\left[\bar{\eta}_I(y_I, \bar{y}_I, u_I)\right] + \|L_I\|\bar{\epsilon}_{\xi_I}(t) + \bar{\epsilon}_{\Delta_{I,2}}(t) + |h_p(t)|\bar{x}_{I,d}, \tag{4.58}$$

where $\bar{x}_{I,d}$ is a bounding estimate of $x_I(0)$ such that $\|x_I(0)\| \le \bar{x}_{I,d}$ and $\bar{H}_p(s)$ is a transfer function with impulse response $\bar{h}_p(t)$ that satisfies $|h_p(t)| \le \bar{h}_p(t)$ for all $t \ge 0$ and can be determined by using a method of Section 3.2.2. Note that, according to Lemma 3.3, $|h_p(t)|$ is exponentially decaying and thus, a conservative estimate $\bar{x}_{I,d}$ affects the detection threshold only during the initial transient period.

So, by using (4.57) and the bounding function $\bar{\chi}_{I,2}(t)$ we obtain $|r_I^{(k)}(t)| \le \bar{r}_{I,2}^{(k)}(t)$, where the detection threshold $\bar{r}_{I,2}^{(k)}(t)$ is given by

$$\bar{r}_{I,2}^{(k)}(t) \triangleq \int_0^t \nu_I^{(k)} e^{-\lambda_I^{(k)}(t-\tau)}\bar{\chi}_{I,2}(\tau)\,\mathrm{d}\tau + \bar{\epsilon}_{\xi_I}(t) \tag{4.59}$$

and can be implemented using simple linear filtering techniques:

$$\bar{r}_{I,2}^{(k)}(t) = \frac{\nu_I^{(k)}}{s + \lambda_I^{(k)}}\left[\bar{\chi}_{I,2}(t)\right] + \bar{\epsilon}_{\xi_I}. \tag{4.60}$$

Figure 4.3 illustrates the implementation of the local filtered fault detection scheme for the I-th subsystem resulting from equations (4.6), (4.49), (4.9) and (4.60).

Christodoulos Keliris

Following the previous analysis, in the absence of any faults the residual signal $r_I^{(k)}(t)$ given in (4.9) is always bounded by the detection threshold $\bar{r}_{I,2}^{(k)}(t)$ given by (4.59). Therefore, this guarantees that there will be no false alarms.

Lemma 4.5. *Consider a distributed system made of N subsystems Σ_I given by (4.47), (4.48). In the absence of any faults, the residual signals $r_I^{(k)}(t)$ given by (4.9) using a filter $H_p(s)$ given by (4.6), where the signals $\hat{y}_I^{(k)}$ are given by (4.49) and (4.50), are bounded by the detection thresholds $\bar{r}_{I,2}^{(k)}(t)$, given by (4.59), thus guaranteeing no false alarms.*

In the following, a fault detectability condition for the special case will be presented.

Theorem 4.3. *Consider the nonlinear interconnected system (4.47), (4.48) with the distributed fault detection scheme described in* (4.6)*, (4.9), (4.49), (4.50), (4.58) and (4.59). A fault in the I-th subsystem occurring at $t = T_0$ is detectable if the filtered fault function $\phi_{I,f}\big(x(t), u_I(t), t\big) \triangleq H_p(s)[\beta_I(t - T_0)\phi_I(x(t), u_I(t))]$ satisfies the following inequality at some time $t > T_0$, for some $k = 1, 2, \ldots, p_I$:*

$$\left| \int_{T_0}^{t} C_I^{(k)} e^{A_{I,0}(t-\tau)} \phi_{I,f}\big(x(\tau), u_I(\tau), \tau\big)\, \mathrm{d}\tau \right| > 2\bar{r}_{I,2}^{(k)}(t). \tag{4.61}$$

Proof. In the presence of a fault that occurs at $t = T_0$, equation (4.54) becomes:

$$\tilde{x}_{I,f}(t) = e^{A_{I,0}t}\tilde{x}_{I,f}(0) + \int_0^t e^{A_{I,0}(t-\tau)}\big(\chi_{I,2}(\tau) + \phi_{I,f}\big(x(\tau), u_I(\tau), \tau\big)\big)\, \mathrm{d}\tau$$

and the residual from (4.55) becomes

$$r_I(t) = \int_0^t C_I e^{A_{I,0}(t-\tau)}\big(\chi_{I,2}(\tau) + \phi_{I,f}\big(x(\tau), u_I(\tau), \tau\big)\big)\, \mathrm{d}\tau + \epsilon_{\xi_I}(t).$$

Note that the term $C_I e^{A_{I,0}t}\tilde{x}_{I,f}(0)$ is omitted since $\tilde{x}_{I,f}(0) = 0$ due to the filters's zero initial conditions. By using the triangle inequality, the k-th element of $r_I(t)$ for $t > T_0$ satisfies:

$$\begin{aligned}
|r_I^{(k)}(t)| &\geq -\left| \int_0^t C_I^{(k)} e^{A_{I,0}(t-\tau)} \chi_{I,2}(\tau)\, \mathrm{d}\tau \right| - |\epsilon_{\xi_I}^{(k)}(t)| \\
&\quad + \left| \int_0^t C_I^{(k)} e^{A_{I,0}(t-\tau)} \phi_{I,f}\big(x(\tau), u_I(\tau), \tau\big)\, \mathrm{d}\tau \right| \\
&\geq -\int_0^t \|C_I^{(k)} e^{A_{I,0}(t-\tau)}\| \|\chi_{I,2}(\tau)\|\, \mathrm{d}\tau - \|\epsilon_{\xi_I}(t)\| \\
&\quad + \left| \int_0^t C_I^{(k)} e^{A_{I,0}(t-\tau)} \phi_{I,f}\big(x(\tau), u_I(\tau), \tau\big)\, \mathrm{d}\tau \right| \\
&\geq -\int_0^t \nu_I^{(k)} e^{-\lambda_I^{(k)}(t-\tau)} \bar{\chi}_{I,2}(\tau)\, \mathrm{d}\tau - \bar{\epsilon}_{\xi_I}(t) \\
&\quad + \left| \int_0^t C_I^{(k)} e^{A_{I,0}(t-\tau)} \phi_{I,f}\big(x(\tau), u_I(\tau), \tau\big)\, \mathrm{d}\tau \right| \\
&\geq -\bar{r}_{I,2}^{(k)}(t) + \left| \int_{T_0}^t C_I^{(k)} e^{A_{I,0}(t-\tau)} \phi_{I,f}\big(x(\tau), u_I(\tau), \tau\big)\, \mathrm{d}\tau \right|.
\end{aligned}$$

Christodoulos Keliris

For fault detection, the inequality $|r_I^{(k)}(t)| > \bar{r}_{I,2}^{(k)}(t)$ must hold at some time $t > T_0$ for some $k = 1, 2, \ldots, p_I$, so the final fault detectability condition given in (4.61) is obtained. □

4.6 Simulation Results

In this section, we consider a numerical example to illustrate some of the concepts developed in this chapter. The example is based on a system of two interconnected one-link manipulators with revolute joints actuated by a DC motor, where the elasticity of the joint can be modeled by a linear tensional spring [68, 77]. The state variable $x_I^{(1)}$ represents the motor position, $x_I^{(2)}$ the motor velocity, $x_I^{(3)}$ the link position and $x_I^{(4)}$ the link velocity. The system dynamics of the two subsystems $I = 1, 2$, with $x_I = \left[x_I^{(1)} x_I^{(2)} x_I^{(3)} x_I^{(4)}\right]$, are given by

$$\dot{x}_I = A_I x_I + f_I + \eta_I$$
$$y_I = C_I x_I + \xi_I$$

where for the first subsystem:

$$A_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -48.6 & -1.25 & 48.6 & 0 \\ 0 & 0 & 0 & 1 \\ 19.5 & 0 & -19.5 & 0 \end{bmatrix}, \eta_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0.1sin(x_1^{(3)}) \end{bmatrix},$$
$$f_1 = \begin{bmatrix} 0 \\ 43.2u \\ 0 \\ -3.33sin(x_1^{(3)}) + 2sin(x_2^{(4)}) \end{bmatrix},$$

and for the second subsystem:

$$A_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -24.3 & -0.625 & 24.3 & 0 \\ 0 & 0 & 0 & 1 \\ 9.75 & 0 & -9.75 & 0 \end{bmatrix}, \eta_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0.1sin(x_2^{(3)}) \end{bmatrix},$$
$$f_2 = \begin{bmatrix} 0 \\ 21.6u \\ 0 \\ -1.665sin(x_2^{(3)}) + 2sin(x_1^{(4)}) \end{bmatrix}.$$

Christodoulos Keliris

We consider two cases for the matrices C_I so we can demonstrate the effectiveness of the approaches given in the chapter; in the first case we use the nonlinear observer design according to Sections 4.2, 4.3 and in the second case we address the specific class of systems described in Section 4.5. In the first case, for both subsystems $I = 1, 2$, the matrix C_I is given by $C_I = [1\ 0\ 0\ 0; 0\ 1\ 0\ 0; 0\ 0\ 0\ 1]$. The input u_I for both subsystems is a sinusoid of magnitude 1 and frequency 1 Hz. In this simple example, we consider an abrupt multiplicative actuator fault in subsystem 1 where the input changes from $u_1 = \bar{u}_1$ to $u_1 = (1 + \psi_1)\bar{u}_1$ for some parameter $\psi_1 \in [-1, 0]$ characterizing the magnitude of the fault. In the simulation example, the fault occurs at $T_0 = 2$ sec with a magnitude $\psi_1 = -0.2$. In addition, the measurement noise is of Gaussian-type with mean $\mu_\xi = 0$, variance $\sigma_\xi^2 = 0.00025$ and maximum magnitude 0.05.

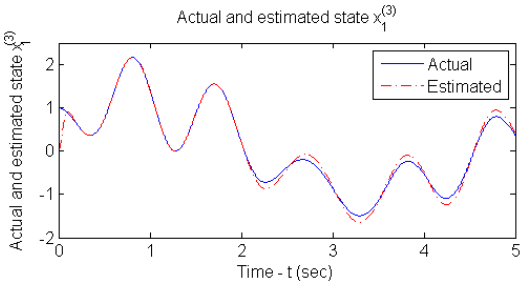
For the first case, we proceed with the nonlinear observer design according to Section 4.2. First, we obtain the "new" A_I and f_I of the nominal system through linearization at the origin of the I-th nonlinear subsystem (so that f_I does not contain first order terms with regards to x_I). Hence, for the first subsystem: $A_1 = [0\ 1\ 0\ 0$; -48.6 -1.25 48.6 0 ; 0 0 0 1 ; 19.5 0 -19.5-3.33 0], $f_1 = [0; 43.2u; 0; -3.33sin(x_1^{(3)}) + 3.33x_1^{(3)} + 2sin(x_2^{(4)})]$ and for the second subsystem: $A_2 = [0\ 1\ 0\ 0$; -24.3 -0.625 24.3 0 ; 0 0 0 1 ; 9.75 0 -9.75-1.665 0], $f_2 = [0; 21.6u; 0; -1.665sin(x_2^{(3)}) + 1.665x_2^{(3)} + 2sin(x_1^{(4)})]$. Just to clarify that, the "new" nonlinear function f_1 does not contain a linear term $x_1^{(3)}$. This is because the Maclaurin series of the term $-3.33sin(x_1^{(3)})$ contains a $-3.33x_1^{(3)}$ term which cancels out $3.33x_1^{(3)}$. Now, the function f_1 satisfies the Lipschitz condition (4.4) with $G_{f_1} = [0\ 0\ 0\ 0;0\ 0\ 0\ 0;0\ 0\ 0\ 0;0\ 0\ 6.66\ 0]$. The nonlinear observer for subsystem 1 is designed so that $\dot{V}_1(\tilde{x}_1) < -\mu_1 V_1(\tilde{x}_1)$ with $\mu_1 = 30$. The specific choice for μ_1 is made because it is required that ρ_1 (which is selected based on (4.26)) satisfies $\rho_1 = \mu_1$ for comparison purposes between the filtering and non-filtering case. To obtain the solution of (4.17) we solve its modified version as an ARE:

$$\bar{A}_1^\top P_1 + P_1 \bar{A}_1 + \epsilon_I G_{f_1}^\top G_{f_1} + \frac{1}{\epsilon_1} P_1 P_1 - \delta_1^2 C_1^\top C_1 = -\epsilon_1' I. \tag{4.62}$$

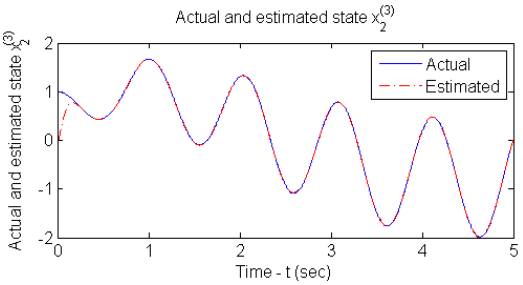
For given $\mu_1 = 30$ and $\epsilon_1' = 0.01$ (generally a small positive scalar), the solution P_1 of (4.62) is obtained and then the observer gain matrix L_1 is calculated from Theorem 4.2. The matrices P_1 and L_1 are found as:

$$P_1 = \begin{bmatrix} 22.9378 & 7.1049 & -11.8737 & -2.4679 \\ 7.1049 & 12.2683 & -8.4383 & 1.0985 \\ -11.8737 & -8.4383 & 15.7520 & 3.7469 \\ -2.4679 & 1.0985 & 3.7469 & 13.6559 \end{bmatrix},$$

Christodoulos Kellis



(a) Subsystem 1.



(b) Subsystem 2.

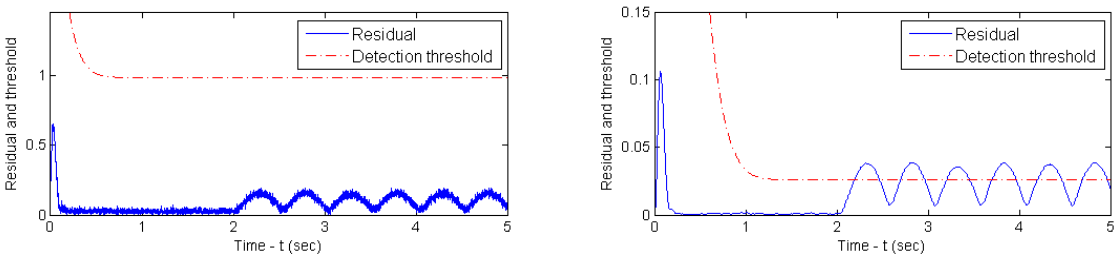
Figure 4.4: Actual and observer estimated link angular position $x_I^{(3)}$, $I = 1, 2$.

$$L_1 = \begin{bmatrix} 22.4565 & -2.1115 & -0.1133 \\ -2.1115 & 44.8900 & -10.8627 \\ 15.8233 & 25.0395 & -12.3573 \\ -0.1133 & -10.8627 & 27.1277 \end{bmatrix}.$$

The eigenvalues of $(A_1 - L_1 C_1)$ are $-25.91 \pm j31.81$ and $-21.95 \pm j1.74$, and therefore they are well damped and they guarantee fast convergence to the actual states for the nominal system. Figure 4.4 illustrates the actual and observer estimated link angular position $x_I^{(3)}$, where for the actual system the initial condition $x_I^{(3)} = 1$ while the observers start from $\hat{x}_I^{(3)} = 0$ for both subsystems $I = 1, 2$. The rest system initial conditions and observer estimates are considered zero. The discrepancy between the actual state and its estimate for subsystem $I = 1$ after $t = 2$ sec is due to the occurrence of the fault.

At first we proceed and implement the residual and threshold signals without the use of filtering as described in Remark 4.3. The residual is implemented according to $r_1(t) = y_1(t) - \hat{y}_1(t)$ and the threshold according to (4.43). The bounds used in this case are $\xi_{1,b} = 0.07$ and $\overline{\Delta f}_{1,1} = 0.1$. In addition the bound on the modeling uncertainty is $\bar{\eta}_1 = 0.1$ and the bound on the system initial conditions is $\bar{x}_{1,d} = 1$. The parameters used for the threshold are $\mu_1 = 30$ and $\theta_1 = 20$. The results are shown in Figure 4.5a where it is shown that the threshold is too conservative thus the fault is not detected. Next, we proceed and implement the proposed fault filtering detection scheme which allows less conservative thresholds to be obtained. The local FDI module that monitors the first subsystem computes the residual (4.9) by using a low-pass filter with transfer function $H_p(s) = \frac{25^2}{(s+50)^2}$ and the fault detection threshold (4.41) with $\rho_1 = 30$, $\gamma_1 = 20$. For comparison purposes, note that these parameters have the same values as in the no filtering case presented before ($\mu_1 = 30$ and $\theta_1 = 20$). For the calculation of the detection threshold, equations (4.40), (4.37) are used with $\bar{\epsilon}_{\xi_1} = 0.002$ and $\bar{\epsilon}_{\Delta_{1,1}} = 0.002$. Note that the transfer function $H_p(s)$ has a non-negative

Christodoulos Keliris



(a) Without filtering.

(b) With filtering.

Figure 4.5: Residual and detection threshold signals based on the nonlinear observer performance.

impulse response $h_p(t)$ and therefore the calculation of $\bar{h}_p(t)$ is not needed, but $h_p(t)$ is used instead. Figure 4.5b shows that the fault is successfully detected at around $t = 2.19$ sec, indicating the effectiveness of the proposed scheme.

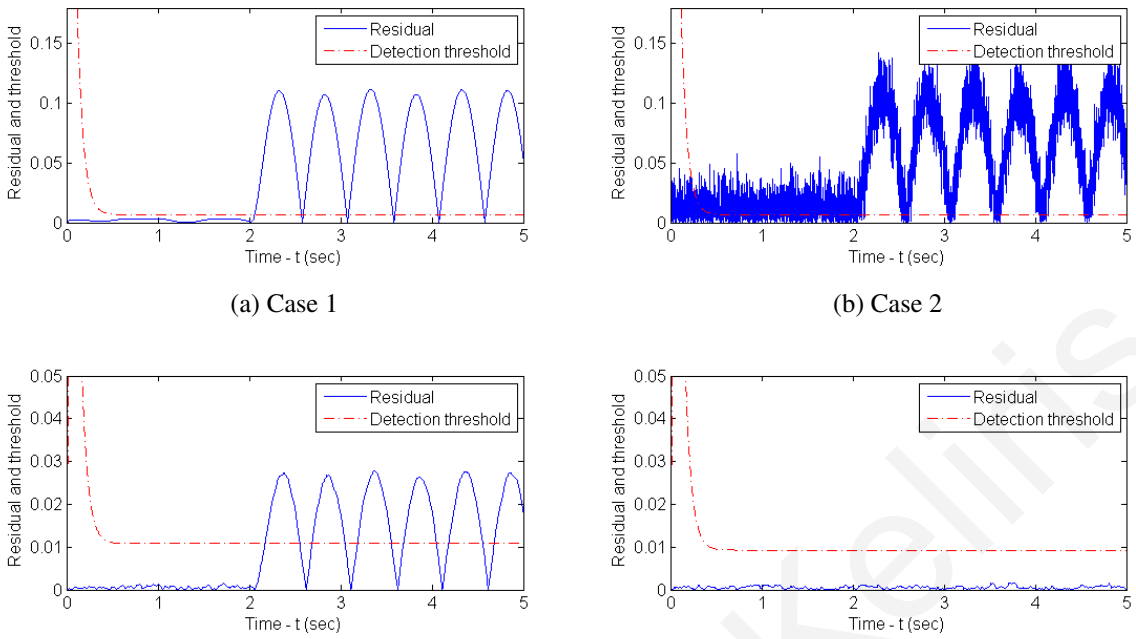
Next, we proceed to demonstrate the example in the special case. For this case we consider that the measurable state variables of the two subsystems are different than in the previous example and now they are identified in both subsystems by the matrices $C_I = [0\ 1\ 0\ 0;\ 0\ 0\ 1\ 0;\ 0\ 0\ 0\ 1]$, so that the special case problem formulation now holds. In this case, the design of a nonlinear observer is not needed for fault detection purposes but the proposed methodology in Section 4.5 is used. Two FDI agents are implemented and each one monitors its associated subsystem. For every measurable state variable, the residual signal is generated according to (4.9) and the corresponding detection threshold using (4.59), (4.58). The estimator $\hat{x}_1(t)$ is implemented according to (4.49) and the gain matrix L_1 is designed so that $(A_1 - L_1C_1)$ is Hurwitz. Similarly, the estimator $\hat{x}_2(t)$ is computed in the second FDI agent, where the gain matrix L_2 is designed so that $(A_2 - L_2C_2)$ is Hurwitz. More specifically, L_1 and L_2 are given by

$$L_1 = \begin{bmatrix} -9.3178 & 0 & 5.0575 \\ 44.6928 & 48.6000 & -10.0984 \\ 0 & 15.0000 & 1.0000 \\ -8.4207 & -19.5000 & 29.0572 \end{bmatrix}, \quad L_2 = \begin{bmatrix} -19.9077 & 0 & 9.4347 \\ 45.5585 & 24.3000 & -9.4874 \\ 0 & 15.0000 & 1.0000 \\ -8.5264 & -9.7500 & 28.8165 \end{bmatrix}.$$

The filter and the bounds that are used are the same as in the previous case for both subsystems and in addition $\bar{\epsilon}_{\Delta_{1,2}} = 0.005$.

In "Case 1" we consider the ideal case where no measurement noise is present. For this case, we omit the use of filtering and therefore the residual and threshold signals are given by modifying accordingly (4.9), (4.59) and (4.58). More specifically, the residual is given

Christodoulos Keliris



(a) Case 1

(b) Case 2

(c) Case 3a

(d) Case 3b

Figure 4.6: Special case: Residual signal and fault detection threshold for Cases 1-3, under different assumptions.

by $r_I(t) = y_I(t) - \hat{y}_I(t)$ and the threshold by $\bar{r}_I^{(k)}(t) = \frac{\nu_I^{(k)}}{s+\lambda_I^{(k)}}\left[\bar{\eta}_I(y_I(t), \bar{y}_I(t), u_I(t))\right] + \nu_I^{(k)} e^{-\lambda_I^{(k)} t}\bar{x}_{I,d}$. For the special case, the simulation results concern the subsystem $I = 1$ and more specifically the residual $r_1^{(4)}$ and threshold $\bar{r}_1^{(4)}$. Figure 4.6a shows the residual and threshold signals in "Case 1" where it can be seen that the fault is successfully detected with no false alarms prior to the fault occurrence.

Next, in "Case 2" the noise term is added to demonstrate its effect in the detectability scheme presented in "Case 1". More specifically, in "Case 2" we use the same residual and threshold signals as in "Case 1", but in the presence of measurement noise. Figure 4.6b demonstrates the results in "Case 2" where it can be seen that the residual exceeds significantly the threshold prior to the fault occurrence, causing false alarms.

Finally, in "Case 3" we demonstrate the effectiveness of the proposed fault filtering approach in the special case where a low-pass filter with transfer function $H_p(s) = \frac{25^2}{(s+50)^2}$ is used for the generation of the residual and threshold signals. As previously the calculation of $\bar{h}_p(t)$ is not needed and $h_p(t)$ is used since it is non-negative. Figure 4.6c shows the residual $r_1^{(4)}$ and detection threshold $\bar{r}_1^{(4)}$ where it is clearly seen that the fault is successfully detected at around $t = 2.15$ sec without any false alarms prior to the fault occurrence. Figure 4.6d shows the results for the subsystem $I = 2$ and more specifically the residual $r_2^{(4)}$ and

Christodoulos Keliris

threshold $\bar{r}_2^{(4)}$. The filter that is used has the same transfer function as before. The simulation results show that the fault effects are not propagated to the second subsystem and the residual does not exceed its threshold, thus the fault is not detected. Therefore, partial fault isolation by identifying the subsystem that the fault has occurred is also achieved and further action to identify the type of fault can be initiated.

4.7 Conclusion

In this chapter a distributed fault detection filtering approach for a class of input-output interconnected, continuous-time, nonlinear systems with modeling uncertainties and measurement noise is presented. Utilizing nonlinear observer design and under certain assumptions, a distributed fault detection scheme is proposed which is inherently tied with the observer characteristics. Moreover, a special case is investigated where the nonlinear observer design is not required for fault detection purposes. In both cases, the respective fault detectability conditions are obtained which characterize in an implicit way a class of detectable faults. The main contribution of this chapter is the novel use of filters as a linear state transformation which allows a more general class of filters to be used and, the exploitation of the proposed nonlinear observer design for the fault detection task, which correlates the fault detection with the observers' performance. Intuitively, filtering allows dampening of the noise and its effects, thus facilitating the design of less conservative thresholds, which enhances fault detectability. The implementation of the scheme relies on a distributed framework, where each subsystem is monitored by a local fault detection agent which requires the input and output measurements of the subsystem that is monitoring and also the measurements of all interconnecting subsystems that are influencing the subsystem that the specific agent is monitoring.

Christodoulos Kallis

Chapter 5

Fault Detection Filtering: Discrete-Time Case

In this chapter, a distributed fault filtering detection scheme is presented for interconnected discrete-time nonlinear systems with modeling uncertainty and in the presence of measurement noise. The proposed filtering scheme is based on the fault diagnosis formulation presented in Chapter 3, where suitable residual and adaptive threshold signals are derived in a continuous-time framework through the use of a certain class of filters and taking advantage of their noise suppression properties. The main contribution of this work is the development of a discrete-time formulation for the design and analysis of a distributed fault filtering approach for interconnected nonlinear systems, which allows for direct implementation in real-world applications as required by digital computers and microprocessors. The development of a discrete time framework leads to certain challenges related to the selection of suitable FIR and IIR filters to be used in the design of the adaptive detection threshold.

The chapter is organized as follows: the problem formulation is given in Section 5.1 and in Section 5.2 the detailed design of the distributed fault detection scheme based on a filtering approach is presented along with some practical issues and proposed solutions. In Section 5.3 the detectability condition that characterizes the class of detectable faults is derived. Finally, in Section 5.4 a simulation example demonstrating the effectiveness of the scheme is presented and finally Section 5.5 provides some concluding remarks.

Christodoulos Keliris

5.1 Problem formulation

We consider an interconnected nonlinear dynamic system comprised of N subsystems Σ_I, $I \in \{1, ..., N\}$. The discrete-time dynamics of each subsystem is described by:

$$x_I(t+1) = f_I(x_I(t), u_I(t)) + g_I(x_I(t), \bar{x}_I(t), u_I(t)) + \eta_I(x_I(t), \bar{x}_I(t), u_I(t), t) + \beta_I(t - T_0)\phi_I(x(t), u_I(t)) \quad (5.1)$$

$$y_I(t) = x_I(t) + \xi_I(t). \quad (5.2)$$

where $t \in \mathbb{N}$ is the discrete time instant, $x_I \in \mathbb{R}^{n_I}$, $u_I \in \mathbb{R}^{m_I}$ and $y_I \in \mathbb{R}^{n_I}$ are the state, input and measured output vectors of the I-th subsystem, respectively, and $x \triangleq \left[x_1^\top, x_2^\top, \ldots, x_N^\top\right]^\top \in \mathbb{R}^n$ is the state vector of the overall system. The vector $\bar{x}_I \in \mathbb{R}^{\bar{n}_I}$ denotes the state variables of neighboring subsystems affecting the I-th subsystem. The function $f_I : \mathbb{R}^{n_I} \times \mathbb{R}^{m_I} \mapsto \mathbb{R}^{n_I}$ is the known local (nominal) function dynamics of the I-th subsystem and $g_I : \mathbb{R}^{n_I} \times \mathbb{R}^{\bar{n}_I} \times \mathbb{R}^{m_I} \mapsto \mathbb{R}^{n_I}$ is the known part of the interconnection function between the I-th and "neighboring" subsystems. The vector function $\eta_I : \mathbb{R}^{n_I} \times \mathbb{R}^{\bar{n}_I} \times \mathbb{R}^{m_I} \times \mathbb{R}^+ \mapsto \mathbb{R}^{n_I}$ denotes the modeling uncertainty associated with the nominal dynamics and $\xi_I \in \mathbb{R}^{n_I}$ represents the measurement noise. The term $\beta_I(t - T_0)\phi_I(x, u_I)$ characterizes the fault function dynamics affecting the I-th subsystem, including its time evolution. More specifically, the term $\phi_I : \mathbb{R}^n \times \mathbb{R}^{m_I} \mapsto \mathbb{R}^{n_I}$ represents the unknown fault function and the term $\beta_I(t-T_0) : \mathbb{R} \mapsto \mathbb{R}^+$ models the time evolution of the fault, where T_0 is the unknown time of the fault occurrence. Note that the fault function ϕ_I may depend on the global state variable vector x and not only on the local state vector x_I. From a practical perspective, this allows faults that are functions of the overall state vector, not only states that are available to the I-th subsystem. In addition, this allows for propagative fault effects to be transferred across neighboring subsystems (as it is the case in real networks such as electric power systems, transportation systems, etc). In this chapter, no particular modeling is considered for the time profile $\beta_I(t - T_0)$ which can be used to model both abrupt and incipient faults. Instead, we simply consider it to be zero prior to the fault occurrence, i.e. $\beta_I(t - T_0) = 0$, for all $t < T_0$. In fact, the faults can be permanent, temporary or even intermittent. As long as they are sufficiently "large" and last for sufficiently long time, then they can be detected by the proposed scheme.

The objective is to design and analyze a distributed fault detection scheme where, to each subsystem Σ_I, we associate a local fault detection agent $\mathcal{F}_I$, which receives local measurements u_I, y_I and partial information from neighboring fault detection agents $\mathcal{F}_J$. Note that

Christodoulos Keliris

each fault detection agent $\mathcal{F}_I$ is not connected to all other agents, but only to a subset (typically small) of neighboring agents, thus constituting a distributed fault detection scheme. More details regarding the distributed nature of the scheme were given in Section 3.1. It is assumed that there exist feedback controllers $\mathcal{U}_I$ for selecting u_I such that some desired control objectives are achieved. The formulation allows for a generic controller $\mathcal{U}_I$ or for a controller that incorporates an advanced fault accommodation scheme. In this chapter, we do not deal explicitly with the control design problem, but instead consider the fault detection issue in the presence of faults ϕ_I, modeling uncertainties η_I and measurement noise ξ_I. Particular emphasis is placed on addressing issues related to measurement noise, which may affect significantly the effectiveness of the fault detection scheme in practical applications.

The following assumptions are used throughout the chapter:

***Assumption* 5.1.** *For each subsystem Σ_I, $I \in \{1, ..., N\}$, the local state variables $x_I(t)$ and the local inputs $u_I(t)$ belong to a known compact region $\mathcal{D}_{x_I}$ and $\mathcal{D}_{u_I}$ respectively before and after the occurrence of a fault, i.e. $x_I(t) \in \mathcal{D}_{x_I}$, $u_I(t) \in \mathcal{D}_{u_I}$ for all $t \geq 0$ (well-posedness).*

***Assumption* 5.2.** *The modeling uncertainty $\eta_I^{(j)}$ (j denotes the j-th component of η_I) is bounded by a known positive functional $\bar{\eta}_I$, i.e.,*

$$|\eta_I^{(j)}(x_I, \bar{x}_I, u_I, t)| \leq \bar{\eta}_I^{(j)}(y_I, \bar{y}_I, u_I), \quad j = 1, 2, \ldots, n_I \tag{5.3}$$

for all $t \in \mathbb{N}$ and for all $(x_I, \bar{x}_I, u_I) \in \mathcal{D}_I$, where $\bar{y}_I \in \mathbb{R}^{\bar{n}_I}$ is the noisy counterpart of $\bar{x}_I$, i.e. $\bar{y}_I = \bar{x}_I + \bar{\xi}_I$, $\bar{\xi}_I \in \mathbb{R}^{\bar{n}_I}$ and $\bar{\eta}_I(y_I, \bar{y}_I, u_I) \geq 0$ is a known bounding function in some compact region of interest $\mathcal{D}_I = \mathcal{D}_{x_I} \times \mathcal{D}_{\bar{x}_I} \times \mathcal{D}_{u_I} \subset \mathbb{R}^{n_I} \times \mathbb{R}^{\bar{n}_I} \times \mathbb{R}^{m_I}$.

Assumption 5.2 characterizes the class of modeling uncertainties being considered. In practice, the system can be modeled more accurately in certain regions of the state space. Therefore, the fact the bound $\bar{\eta}_I$ is a function of y_I, $\bar{y}_I$ and u_I provides more flexibility by allowing the designer to take into consideration any prior knowledge of the system that may be available.

Similarly to the continuous time case, the notation $y(t) = H(z)\big[x(t)\big]$ denotes the output $y(t)$ of any discrete time signal $x(t)$ which is passed through a filter with dicrete-time transfer function $H(z)$. In terms of more rigorous notation, let $h(t)$ be the discrete time impulse response associated with $H(z)$; i.e. $h(t) \triangleq \mathcal{Z}^{-1}[H(z)]$, where $\mathcal{Z}^{-1}$ is the inverse Z-transform operator. Then $y(t) = H(z)\big[x(t)\big] = \sum_{k=0}^{t} h(k)x(t-k)$.

To dampen the effect of measurement uncertainty $\xi_I(t)$, each measured variable $y_I^{(j)}$ (j-th component of y_I) is filtered by $H(z)$, where $H(z)$ is a p-th order, asymptotically stable

Christodoulos Keliris

(poles lie inside the open unit disc $|z| = 1$) filter with proper transfer function

$$H(z) = \frac{d_0 + d_1 z^{-1} + d_2 z^{-2} + \ldots + d_p z^{-p}}{1 + c_1 z^{-1} + \ldots + c_p z^{-p}}. \tag{5.4}$$

Generally, each measured variable $y_I^{(j)}(t)$ can be filtered by a different filter but in this chapter, without loss of generality, we consider $H(z)$ to be the same for all the output variables in order to simplify the notation and presentation. In addition, note that the form of $H(z)$ allows both IIR and FIR types of digital filters.

The filter $H(z)$ can be written as $H(z) = zH_p(z)$ where $H_p(z)$ is the strictly proper transfer function

$$H_p(z) = \frac{d_0 z^{-1} + d_1 z^{-2} + d_2 z^{-3} + \ldots + d_p z^{-(p+1)}}{1 + c_1 z^{-1} + \ldots + c_p z^{-p}}. \tag{5.5}$$

Note that the filter $H_p(z)$ is also asymptotically stable since it comprises of the same poles as $H(z)$ with an additional pole at $z = 0$ (inside $|z| = 1$). Since the filters $H(z)$ and $H_p(z)$ (with impulse responses $h(t)$ and $h_p(t)$ respectively) are asymptotically stable, they are also BIBO stable. Therefore, for bounded measurement noise $\xi_I(t)$, the filtered measurement noise $\epsilon_{\xi_I}(t) \triangleq H(z)\left[\xi_I(t)\right]$ is bounded as follows:

$$|\epsilon_{\xi_I}^{(j)}(t)| \leq \bar{\epsilon}_{\xi_I}^{(j)}(t) \quad j = 1, 2, \ldots, n_I, \tag{5.6}$$

where $\bar{\epsilon}_{\xi_I}^{(j)}$ are computable bounding functions. Depending on the noise characteristics, $H(z)$ can be selected to reduce the bounding function $\bar{\epsilon}_{\xi_I}^{(j)}(t)$.

5.2 Distributed Fault Detection

5.2.1 Residual Signal Generation

By locally filtering the output signal $y_I(t)$ we obtain the filtered output $w_I(t)$:

$$\begin{aligned} w_I(t) &= H(z)\left[y_I(t)\right] \\ &= H(z)\left[x_I(t) + \xi_I(t)\right]. \end{aligned} \tag{5.7}$$

By using $\epsilon_{\xi_I}(t) = H(z)\left[\xi_I(t)\right]$ and $H(z) = zH_p(z)$ we obtain:

$$w_I(t) = H_p(z)\left[z\left[x_I(t)\right]\right] + \epsilon_{\xi_I}(t). \tag{5.8}$$

Using the time-shift property of the Z-transform we have:

$$z\big[x_I(t)\big] = x_I(t+1) + z\big[x_I(0)\delta(t)\big], \tag{5.9}$$

Christodoulos Kellinis

where $\delta(t)$ is the discrete delta function.

Using (5.9) and (5.1), (5.8) becomes

$$\begin{aligned} w_I(t) &= H_p(z)\big[x_I(t+1) + z\big[x_I(0)\delta(t)\big]\big] + \epsilon_{\xi_I}(t) \\ &= H_p(z)\big[x_I(t+1)\big] + H(z)\big[x_I(0)\delta(t)\big] + \epsilon_{\xi_I}(t) \\ &= H_p(z)\big[f_I(x_I(t), u_I(t)) + g_I(x_I(t), \bar{x}_I(t), u_I(t)) + \eta_I(x_I(t), \bar{x}_I(t), u_I(t), t) \\ &\qquad + \beta_I(t - T_0)\phi_I(x(t), u_I(t))\big] + h(t)x_I(0) + \epsilon_{\xi_I}(t). \end{aligned} \tag{5.10}$$

Based on (5.1), a local estimation model *under fault-free operation* is generated by considering the known components and by ignoring the unknown modeling uncertainty η_I, the faults ϕ_I and by replacing x_I and $\bar{x}_I$ with y_I and $\bar{y}_I$ respectively:

$$\hat{x}_I(t+1) = f_I\big(y_I(t), u_I(t)\big) + g_I\big(y_I(t), \bar{y}_I(t), u_I(t)\big) \tag{5.11}$$

$$\hat{w}_I(t) = H(z)\left[\hat{x}_I(t)\right] \tag{5.12}$$

with initial condition $\hat{x}_I(0) = y_I(0)$.

The local estimation model described by (5.11), (5.12) is used to generate the residual signal that will be utilized for fault detection purposes. Specifically, the local residual error $r_I(t)$ to be used for fault detection is defined as:

$$r_I(t) \triangleq w_I(t) - \hat{w}_I(t). \tag{5.13}$$

This residual constitutes the basis of the fault detection scheme and it is readily computable from equations (5.7), (5.11) and (5.12). A fault in the overall system is said to be detectable (i.e. a detection decision can be made) when $|r_I^{(j)}(t)| > \bar{r}_I^{(j)}(t)$, for at least one component j in any local subsystem Σ_I, where $\bar{r}_I^{(j)}(t)$ is the detection threshold (to be specified later on).

5.2.2 Detection Threshold

To derive a suitable detection threshold, we consider the maximum effect of the uncertainty on the residual signal, in the absence of any faults. Using a similar procedure as in the derivation of (5.10), $\hat{w}_I(t)$ satisfies:

$$\hat{w}_I(t) = H_p(z)\big[f_I\big(y_I(t), u_I(t)\big) + g_I\big(y_I(t), \bar{y}_I(t), u_I(t)\big)\big] + h(t)y_I(0). \tag{5.14}$$

Prior to the fault ($t < T_0$), the local residual error can be written using equations (5.10), (5.14) and (5.13) as:

$$r_I(t) = H_p(z)[\chi_I(t)] + \epsilon_{\xi_I}(t) - \xi_I(0)h(t), \tag{5.15}$$

Christodoulos Keliris

where the total uncertainty term $\chi_I(t)$ is defined as:

$$\chi_I(t) \triangleq \Delta f_I(t) + \Delta g_I(t) + \eta_I\big(x_I(t), \bar{x}_I(t), u_I(t), t\big), \tag{5.16}$$

with

$$\Delta f_I(t) \triangleq f_I\big(x_I(t), u_I(t)\big) - f_I\big(y_I(t), u_I(t)\big), \tag{5.17}$$

$$\Delta g_I(t) \triangleq g_I\big(x_I(t), \bar{x}_I(t), u_I(t)\big) - g_I\big(y_I(t), \bar{y}_I(t), u_I(t)\big). \tag{5.18}$$

By taking the absolute value of (5.15) element-wise and using the triangle inequality we obtain:

$$\begin{aligned} |r_I^{(j)}(t)| &\leq |H_p(z)[\chi_I^{(j)}(t)]| + |\epsilon_{\xi_I}^{(j)}(t)| + |\xi_I^{(j)}(0)h(t)| \\ &\leq \sum_{k=0}^{t} |h_p(k)||\chi_I^{(j)}(t-k)| + \bar{\epsilon}_{\xi_I}^{(j)}(t) + |\xi_I^{(j)}(0)||h(t)|. \end{aligned}$$

Based on the above inequality, the adaptive detection threshold $\bar{r}_I^{(j)}(t)$ is set to:

$$\bar{r}_I^{(j)}(t) \triangleq \bar{H}_p(z)[\bar{\chi}_I^{(j)}(t)] + \bar{\epsilon}_{\xi_I}^{(j)}(t) + \xi_{I,b}^{(j)}|h(t)|, \tag{5.19}$$

where $\bar{H}_p(z)$ is a filter with impulse response $\bar{h}_p(t) \geq |h_p(t)|$, $\bar{\chi}_I^{(j)}(t)$ is the bound on the total uncertainty term $\chi_I^{(j)}(t)$, i.e., $0 \leq |\chi_I^{(j)}(t)| \leq \bar{\chi}_I^{(j)}(t)$ and is given by

$$\bar{\chi}_I^{(j)}(t) \triangleq \bar{\Delta} f_I^{(j)} + \bar{\Delta} g_I^{(j)} + \bar{\eta}_I^{(j)}\big(y_I(t), \bar{y}_I(t), u_I(t)\big), \tag{5.20}$$

where

$$\bar{\Delta} f_I^{(j)} \triangleq \sup_{\substack{(x_I, u_I) \in \mathcal{D}_{x_I} \times \mathcal{D}_{u_I} \\ \xi_I \in \mathcal{D}_{\xi_I}}} |f_I^{(j)}\big(x_I, u_I\big) - f_I^{(j)}\big(x_I + \xi_I, u_I\big)| \tag{5.21}$$

$$\bar{\Delta} g_I^{(j)} \triangleq \sup_{\substack{(x_I, \bar{x}_I, u_I) \in \mathcal{D}_I \\ (\xi_I, \bar{\xi}_I) \in \mathcal{D}_{\xi_I} \times \mathcal{D}_{\bar{\xi}_I}}} |g_I^{(j)}\big(x_I, \bar{x}_I, u_I\big) - g_I^{(j)}\big(x_I + \xi_I, \bar{x}_I + \bar{\xi}_I, u_I\big)| \tag{5.22}$$

and $\xi_{I,b}^{(j)}$ is a bounding estimate of $\xi_I^{(j)}(0)$, i.e. $|\xi_I^{(j)}(0)| \leq \xi_{I,b}^{(j)}$. Note that, even for a conservative bound $\xi_{I,b}^{(j)}$, the term $\xi_{I,b}^{(j)}|h(t)|$ affects the detection threshold only during the initial transient. This is because the impulse response $h(t)$ of a proper and asymptotically stable transfer function $H(z)$ converges to zero exponentially fast. In mathematical terms, let (A, B, C, D) be a state space representation of the transfer function $H(z)$. Then $\|A^t\| \leq m\rho^t$ for all $t \in \mathbb{N}$ for some $\rho \in [0, 1)$ and $m > 0$ [8] thus $h(t)$ satisfies $|h(t)| \leq \|C\|m\rho^{t-1}\|B\|s[t-1] + |D|\delta[t]$

Christodoulos Kellis

where $s[t-1]$ is the discrete step function [7]. Therefore, there exist $\mu > 0$, $\gamma \in [0, 1)$ such that for all $t \in \mathbb{N}$ the following inequality holds:

$$|h(t)| \leq \mu \gamma^t s[t]. \tag{5.23}$$

From a practical viewpoint, the implementation of the threshold $\bar{r}_I$ requires the bounds $\bar{\Delta} f_I$ and $\bar{\Delta} g_I$ given in (5.21) and (5.22) respectively. One approach to derive these bounds is to consider a local Lipschitz condition, i.e. in the case of the function $g_I^{(j)}$:

$$|g_I^{(j)}\big(x_I, \bar{x}_I, u_I\big) - g_I^{(j)}\big(x_I + \xi_I, \bar{x}_I + \bar{\xi}_I, u_I\big)| \leq L_{g_I^{(j)}} \|[\xi_I \quad \bar{\xi}_I]^\top\|, \tag{5.24}$$

where $L_{g_I^{(j)}}$ is the Lipschitz constant for the function $g_I^{(j)}$ with respect to $(x_I, \bar{x}_I)$ in the region $\mathcal{D}_{x_I} \times \mathcal{D}_{\bar{x}_I}$. Therefore, if we have a uniform bound on the measurement noise, i.e. $\|\xi(t)\| \leq \xi_{I,b}$, $\|\bar{\xi}(t)\| \leq \bar{\xi}_{I,b} \quad \forall t \in \mathbb{N}$, then we can derive the bound $\bar{\Delta} g_I$. A similar procedure can be followed for the derivation of the bound $\bar{\Delta} f_I$.

A less conservative detection threshold can be obtained by writing the residual (5.15) as

$$r_I(t) = H_p(z)[\eta_I\big(x_I(t), \bar{x}_I(t), u_I(t), t\big)] + \epsilon_{\Delta_I}(t) + \epsilon_{\xi_I}(t) - \xi_I(0)h(t) \tag{5.25}$$

where

$$\epsilon_{\Delta_I}(t) \triangleq H_p(z)[\Delta f_I(t) + \Delta g_I(t)] \tag{5.26}$$

and by making the following assumption:

***Assumption* 5.3.** *The filtered function mismatch term $\epsilon_{\Delta_I}(t)$ is bounded as follows:*

$$|\epsilon_{\Delta_I}^{(j)}(t)| \leq \bar{\epsilon}_{\Delta_I}^{(j)}(t) \quad j = 1, 2, \ldots, n_I, \tag{5.27}$$

where $\bar{\epsilon}_{\Delta_I}^{(j)}(t)$ is a computable bounding function.

Assumption 5.3 is based on the fact that filtering dampens the error effect of measurement noise evident in the function mismatch term $\Delta f_I^{(j)}(t) + \Delta g_I^{(j)}(t)$. A suitable selection of $\bar{\epsilon}_{\Delta_I}$ can be made through the use of simulations (i.e. Monte Carlo methods) by filtering the function mismatch term $\Delta f_I^{(j)}(t) + \Delta g_I^{(j)}(t)$ using the known nominal and interconnection function dynamics and the available noise characteristics.

Therefore, the detection threshold becomes

$$\bar{r}_I^{(j)}(t) = \bar{H}_p(z)[\bar{\eta}_I^{(j)}\big(y_I(t), \bar{y}_I(t), u_I(t)\big)] + \bar{\epsilon}_{\Delta_I}^{(j)}(t) + \bar{\epsilon}_{\xi_I}^{(j)}(t) + \xi_{I,b}^{(j)}|h(t)|. \tag{5.28}$$

Figure 5.1 illustrates the implementation of the local filtered fault detection scheme for the I-th subsystem resulting from equations (5.4), (5.7), (5.11), (5.12), (5.13) and (5.19).

Christodoulos Keliris

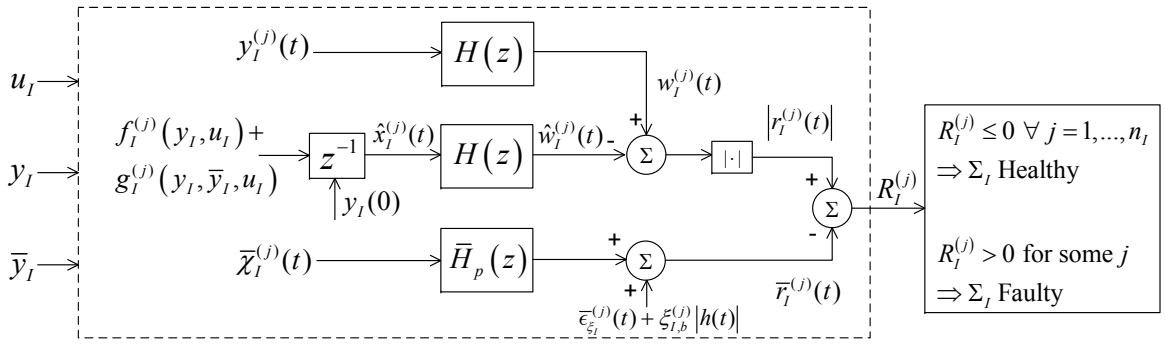


Figure 5.1: Local Filtered Fault Detection Scheme.

Following the previous analysis, in the absence of any faults the residual signal $r_I^{(j)}(t)$ given in (5.13) is always bounded by the detection threshold $\bar{r}_I^{(j)}(t)$ given by (5.19) or (5.28). Therefore, this guarantees that there will be no false alarms. The findings of the preceding analysis are summarized in the following Lemma.

Lemma 5.1. *Consider a distributed system constituted by N subsystems Σ_I given by (5.1), (5.2) with the fault detection scheme described by (5.7), (5.11) and (5.12). Then for any $j = 1, 2, \ldots, n_I$, the residual signal given by (5.13) satisfies*

$$|r_I^{(j)}(t)| \le \bar{r}_I^{(j)}(t), \quad \forall t \in [0, T_0).$$

where the detection threshold $\bar{r}_I^{(j)}(t)$ is given by (5.19) or (5.28).

In general, the distributed fault detection scheme is constituted by N local filtered fault detection modules, one for each subsystem Σ_I. Each subsystem is monitored by a local fault detection module which requires the input and output measurements of the subsystem that is monitoring and also the measurements of all interconnecting subsystems that are influencing the subsystem that the specific module is monitoring. Therefore, there is the need of a communication infrastructure enabling the information exchange between the local fault detection modules depending on their interconnections.

5.2.3 Selection of filter $\bar{H}_p(z)$

A practical issue that requires consideration is the selection of the filter $\bar{H}_p(z)$ whose impulse response must satisfy $|h_p(t)| \le \bar{h}_p(t)$ as stated before. In the case where the impulse response $h_p(t)$ is non-negative, the selection $\bar{H}_p(z) = H_p(z)$ is trivial. Sufficient conditions for non-negative impulse response for a class of discrete-time transfer functions are given in [59]. In the following, we present two methods for choosing $\bar{H}_p(z)$, one considering $H(z)$ as a digital IIR filter and the other one as a FIR filter.

Christodoulos Kallis

First we consider the case where $H(z)$ is an IIR filter. Due to the way $H_p(z)$ was defined, $H_p(z)$ is strictly proper and asymptotically stable. Hence, similarly to (5.23), the impulse response $h_p(t)$ satisfies $|h_p(t)| \leq \kappa\lambda^t$ for all $t \in \mathbb{N}$, for some $\kappa > 0$ and $\lambda \in [0, 1)$. Since $|h_p(t)| \leq \bar{h}_p(t)$ must hold, the impulse response $\bar{h}_p(t)$ can be selected as $\bar{h}_p(t) = \kappa\lambda^t$ and thus $\bar{H}_p(z) = \frac{\kappa}{1-\lambda z^{-1}}$.

Now, let's consider the case in which $H(z)$ is a FIR filter. FIR filters have several advantages, as they are inherently stable and can easily be designed to be linear-phase which corresponds to uniform delay at all frequencies. Let $H(z)$ be a p-th order FIR filter given by $H(z) = \sum_{k=0}^{p} d_k z^{-k}$. Therefore, $H_p(z) = z^{-1}H(z) = \sum_{k=0}^{p} d_k z^{-(k+1)}$ and $\bar{h}_p(t)$ can be selected as $\bar{h}_p(t) = |h_p(t)|$ which leads to the FIR filter $\bar{H}_p(z) = \sum_{k=0}^{p} |d_k| z^{-(k+1)}$.

5.3 Fault detectability analysis

In this section, fault detectability conditions for the aforementioned fault detection scheme are addressed. In order to derive these conditions, we take into consideration the occurrence of a fault ϕ_I at an unknown time $t = T_0$. The fault detectability analysis constitutes a theoretical result that characterizes quantitatively (and in implicit way) the class of faults detectable by the proposed scheme. The following Theorem is based on threshold (5.19), however it can be readily shown that the same result holds in the case where threshold (5.28) is used.

Theorem 5.1 (Fault Detectability)**.** *Consider the nonlinear interconnected system (5.1), (5.2) with the distributed fault detection scheme described in (5.4), (5.7), (5.11), (5.12), (5.13) and (5.19). A fault in the I-th subsystem initiated at $t = T_0$ is detectable if the fault function $\phi_I^{(j)}(x(t), u_I(t))$ satisfies the following inequality at some time $t > T_0$ and for some $j = 1, 2, \ldots, n_I$:*

$$\left| \sum_{k=T_0}^{t} h_p(t-k)\beta_I(k-T_0)\phi_I^{(j)}\big(x(k), u_I(k)\big) \right| > 2\bar{r}_I^{(j)}(t). \tag{5.29}$$

Proof. In the presence of a fault that occurs at $t = T_0$, equation (5.15) becomes:

$$r_I(t) = H_p(z)[\chi_I(t) + \beta_I(t-T_0)\phi_I\big(x(t), u_I(t)\big)] + \epsilon_{\xi_I}(t) - \xi_I(0)h(t). \tag{5.30}$$

By taking the absolute value element-wise and using the triangle inequality, the j-th component of (5.30) satisfies

$$\begin{aligned} |r_I^{(j)}(t)| \geq &- |H_p(z)[\chi_I^{(j)}(t)]| - |\epsilon_{\xi_I}^{(j)}(t)| - |\xi_I^{(j)}(0)h(t)| \\ &+ |H_p(z)[\beta_I(t-T_0)\phi_I^{(j)}\big(x(t), u_I(t)\big)]|. \end{aligned} \tag{5.31}$$

Christodoulos Keliris

and by using a similar procedure as in the derivation of (5.19), (5.31) becomes

$$\begin{aligned}|r_I^{(j)}(t)| \geq & -\bar{H}_p(z)[\bar{\chi}_I^{(j)}(t)] - \bar{\epsilon}_{\xi_I}^{(j)}(t) - \xi_{I,b}^{(j)}(0)|h(t)| \\ & + |H_p(z)[\beta_I(t-T_0)\phi_I^{(j)}\big(x(t),u_I(t)\big)]| \\ \geq & -\bar{r}_I^{(j)}(t) + |H_p(z)[\beta_I(t-T_0)\phi_I^{(j)}\big(x(t),u_I(t)\big)]|. \end{aligned} \tag{5.32}$$

For fault detection, the inequality $|r_I^{(j)}(t)| > \bar{r}_I^{(j)}(t)$ must hold at some time $t > T_0$ and for some $j = 1, 2, \ldots, n_I$, so the final fault detectability condition is obtained:

$$\Big|H_p(z)\big[\beta_I(t-T_0)\phi_I^{(j)}\big(x(t),u_I(t)\big)\big]\Big| > 2\bar{r}_I^{(j)}(t).$$

This can be rewritten in the summation form (5.29) of the Theorem. □

This theorem provides a sufficient condition for the implicit characterization of a class of faults that can be detected by the proposed fault detection scheme. Clearly, the fault functions $\phi_I(x, u_I)$ are typically unknown and therefore this condition cannot be checked apriori. However, it provides useful intuition about the type of faults that are detectable.

5.4 Simulation Results

In this section, we consider a numerical example based on a system of two inverted pendulums connected by a spring. The discrete time models of the two subsystems $I = 1, 2$ are obtained from the continuous time version [49] by using a forward Euler discretization with a time step $T_s = 0.0001$s and are given by

$$\begin{aligned} x_I^{(1)}(t+1) =& x_I^{(1)}(t) + T_s x_I^{(2)}(t) \\ x_I^{(2)}(t+1) =& x_I^{(2)}(t) + T_s\left(f_I^{(2)}(t) + g_I^{(2)}(t) + \eta_I^{(2)}(t)\right) \end{aligned}$$

where for the first subsystem the nominal and interconnection functions are given by:

$$\begin{aligned} f_1^{(2)}(t) =& \left(\frac{m_1 gr}{J_1} - \frac{kr^2}{4J_1}\right)\sin(x_1^{(1)}(t)) + \frac{kr}{2J_1}(l-b) + \frac{u_1}{J_1} \\ g_1^{(2)}(t) =& \frac{kr^2}{4J_1}\sin(x_2^{(2)}(t)) \end{aligned}$$

and for the second subsystem the respective functions are

$$\begin{aligned} f_2^{(2)}(t) =& \left(\frac{m_2 gr}{J_2} - \frac{kr^2}{4J_2}\right)\sin(x_2^{(1)}(t)) + \frac{kr}{2J_2}(l-b) + \frac{u_2}{J_2} \\ g_2^{(2)}(t) =& \frac{kr^2}{4J_2}\sin(x_1^{(2)}(t)). \end{aligned}$$

Christodoulos Keliris

The parameters that are used in the simulation are: $m_1 = 2$kg, $m_2 = 2.5$kg, $J_1 = 0.5$kg, $J_2 = 0.625$kg, $k = 30$N/m, $l = 0.5$m, $b = 0.4$m and $g = 9.81$m/s^2. The nominal functions $f_1^{(2)}$ and $f_2^{(2)}$ satisfy the Lipschitz condition with $L_{f_1^{(2)}} = 33.5$ and $L_{f_2^{(2)}} = 21$, respectively, under healthy mode of operation. The interconnection terms $g_1^{(2)}$ and $g_2^{(2)}$ satisfy the Lipschitz condition with $L_{g_1^{(2)}} = 3.75$ and $L_{g_2^{(2)}} = 3$ respectively. The modeling uncertainties of both subsystems are assumed to be $\eta_1^{(2)} = \eta_2^{(2)} = 0.1\sin(10t)$, thus $\bar{\eta}_1^{(2)} = \bar{\eta}_2^{(2)} = 0.1$. The inputs u_I are derived based on a simple decentralized proportional feedback controller that stabilizes each subsystem and are given by $u_I = 20e_I$, I=1,2 where $e_I = -x_I^{(1)}$ is the tracking error. In this example, we consider an abrupt multiplicative actuator fault in subsystem 1 where the input changes to $u_1 = (1 + \theta_1)\bar{u}_1$, where $\bar{u}_1$ is the nominal control input in the non-fault case and $\theta_1 \in [-1, 0]$ is the parameter characterizing the magnitude of the fault. The fault occurs at $T_0 = 5$ sec with a magnitude $\theta_1 = -0.2$.

The measurement noise ξ_I is implemented as a uniform random number in the range $[-0.05, 0.05]$. For the fault detectability task, the residual is given by (5.13) and the detection threshold is implemented according to (5.19) using (5.20). The proposed fault detection scheme is implemented under two cases, the first using FIR filters for $H(z)$ and the second using IIR filters. In the first case, the filter $H(z)$ is designed as a 10-th order FIR lowpass filter with normalized cutoff frequency 0.2 and utilizing a Hamming window (using the `fir1` command in Matlab). The transfer function of $H(z)$ is given by $H(z) = \sum_{k=0}^{10} d_k z^{-k}$ and explained in Section 5.2.3 the filter $\bar{H}_p(z)$ is given by $\bar{H}_p(z) = \sum_{k=0}^{10} |d_k| z^{-(k+1)}$. Using the aforementioned filter $H(z)$ the bounds on the filtered noise are $\bar{\epsilon}_{\xi_I}^{(j)} = 2\text{e}{-5}$ for $I = 1, 2$ and $j = 1, 2$. In the second case, the filter $H(z)$ is designed as a 5-th order lowpass Butterworth digital filter with the same cutoff frequency as before. In this case, we use $\bar{H}_p(z) = \frac{0.5}{1-0.9z^{-1}}$ for which its impulse response satisfies $|h_p(t)| \leq \bar{h}_p(t)$ as it is shown in Figure 5.2. Using the IIR type $H(z)$ the bounds on the filtered noise are $\bar{\epsilon}_{\xi_I}^{(j)} = 2.5\text{e}{-6}$ for $I = 1, 2$ and $j = 1, 2$.

According to the proposed fault detection scheme, for each measured variable a respec-

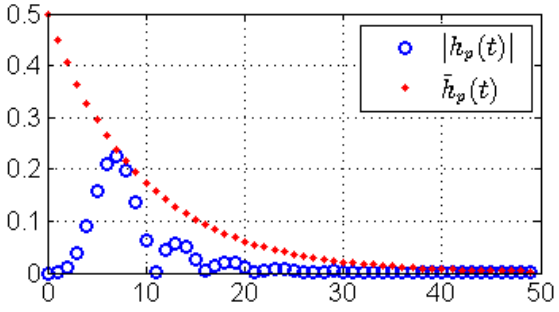
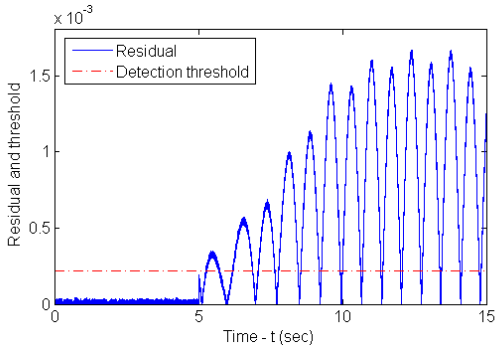
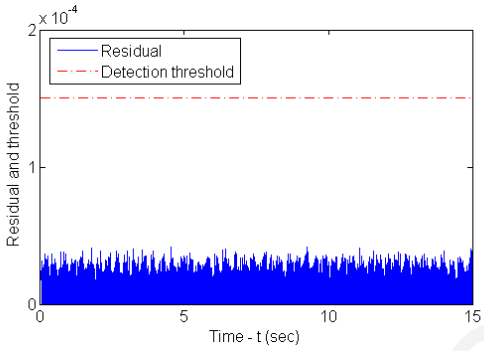


Figure 5.2: Impulse response $|h_p(t)|$ vs $\bar{h}_p(t)$.

Christodoulos Kalis

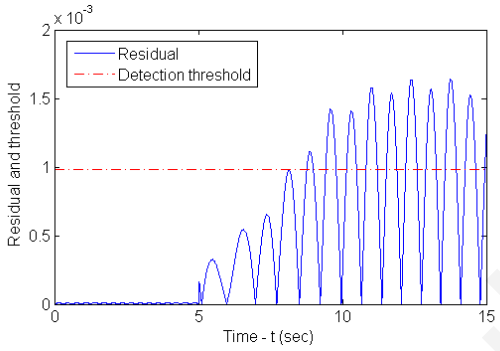


(a) Fault detector monitoring measurement $y_1^{(2)}$.

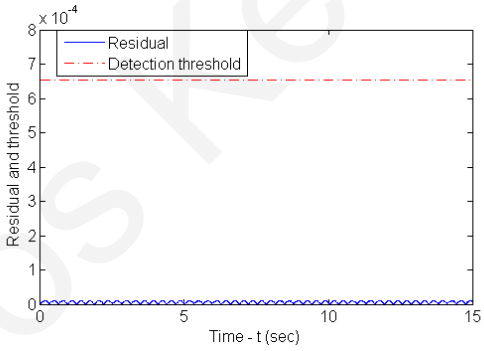


(b) Fault detector monitoring measurement $y_2^{(2)}$.

Figure 5.3: Residual and detection threshold for measurements $y_1^{(2)}$, $y_2^{(2)}$ using FIR filters.



(a) Fault detector monitoring measurement $y_1^{(2)}$.



(b) Fault detector monitoring measurement $y_2^{(2)}$.

Figure 5.4: Residual and detection threshold for measurements $y_1^{(2)}$, $y_2^{(2)}$ using IIR filters.

tive fault detector monitoring it is designed. The simulation results are presented for the fault detection modules that monitor the measurement $y_1^{(2)}$ (first subsystem) and $y_2^{(2)}$ (second subsystem) using FIR and IIR filters in Figures 5.3 and 5.4 respectively. The simulation results that correspond to the detection module that monitors the measurements $y_1^{(2)}$ using FIR filters are shown in Figure 5.3a where it is clearly seen that the residual significantly exceeds its detection threshold after the fault occurs, hence the fault is detected at around $t = 5.3$ s. It must be noted that there are no false alarms prior to the fault occurrence. The corresponding results of the module that monitors $y_2^{(2)}$ are shown in Figure 5.3b where it is seen that the residual signal is always below its threshold signal and therefore no fault is detected. Similarly, the results in the case of the fault detection modules making use of IIR filters, that monitor the measurements $y_1^{(2)}$ and $y_2^{(2)}$ are shown in Figure 5.4a and 5.4b respectively. Again, the fault is detected only by the detection module that monitors $y_1^{(2)}$ at around $t = 8.7$ s. So, the fault is detected at a later stage in comparison to the case where FIR filters are used but as it can be seen from Figure 5.3 the noise is not attenuated as effectively by he FIR filters

Christodoulos Keliris

as in the case where IIR filters are used. In both cases, the fault effects appearing in the first subsystem are not propagated to the second subsystem and therefore no fault is detected by the corresponding fault detection modules. Furthermore, partial fault isolation is achieved in the sense that the faulty subsystem is identified and further fault isolation procedures can be initiated to identify the type of fault that has occurred and aid in fault accommodation.

5.5 Conclusion

In this chapter a distributed fault detection filtering approach for discrete-time, nonlinear systems with modeling uncertainties and measurement noise is presented which constitutes the discrete-time extension of Chapter 3. The scheme makes use of filters in a distributed framework in a way that takes advantage of their noise dampening characteristics and allows the derivation of tight threshold signals enhancing fault detectability.

Christodoulos Kellris

Chapter 6

Distributed Fault Diagnosis for Process and Sensor Faults

The primary objective of this chapter is the derivation of a distributed fault diagnosis approach dealing with process and sensor faults in an integrated way, utilizing a specifically designed scheme that encompasses important characteristics regarding fault propagation among subsystems. Additionally, detectability conditions for process and sensor faults are derived, characterizing quantitatively the class of detectable faults. The scheme is comprised of a set of interacting fault detection agents, in which each subsystem is monitored by its respective detection agent. As it is shown, a process fault that occurs in a subsystem can only be detected by its respective detection agent, whilst a sensor fault that occurs in a subsystem may also be detected by the detection agents of neighboring subsystems that are affected by the subsystem where the sensor fault occurred. This differentiating element is exploited in order to derive a high-level fault isolation scheme, able to provide information regarding the the type and location of the fault that has occurred. Therefore, the proposed distributed fault detection approach encompasses significant benefits for the fault isolation task which can be exploited by a more sophisticated isolation scheme to pinpoint the exact fault that occurred.

The chapter is organized as follows: the problem formulation is given in Section 6.1 and the detailed design of the distributed fault detection scheme is presented in Section 6.2. In Section 6.3, the detectability conditions that characterize the class of detectable process and sensor faults are derived. In Section 6.4, fault propagation issues are investigated, and in Section 6.5, a high-level fault isolation scheme is proposed. In Section 6.6, a simulation example demonstrating the effectiveness of the scheme is presented and, finally, in Section 6.7

Christodoulos Kellis

some concluding remarks are given.

6.1 Problem formulation

We consider an interconnected nonlinear dynamic system comprised of N subsystems Σ_I, $I \in \{1, ..., N\}$. The discrete-time dynamics of each subsystem is described by:

$$\Sigma_I : \begin{cases} x_I(t+1) = A_I x_I(t) + f_I(y_I^0(t), \bar{y}_I^0(t), u_I(t)) + \eta_I(x_I(t), \bar{x}_I(t), u_I(t), t) \\ \qquad\qquad +\beta_I^x(t - T_0^x)\phi_I(x(t), u_I(t)) \qquad (6.1) \\ y_I(t) = C_I x_I(t) + \xi_I(t) + \beta_I^y(t - T_0^y)\theta_I(t) \qquad (6.2) \end{cases}$$

where $t \in \mathbb{N}$ is the discrete time instant, $x_I \in \mathbb{R}^{n_I}$, $u_I \in \mathbb{R}^{m_I}$ and $y_I \in \mathbb{R}^{p_I}$ are the state, input and measured output vectors of the I-th subsystem, respectively, and $x \triangleq \left[x_1^\top, x_2^\top, \ldots, x_N^\top\right]^\top \in \mathbb{R}^n$ is the state vector of the overall system. Note that the distributed fault diagnosis scheme (to be presented) is composed by N local fault detection agents $\mathcal{F}_I$, $I \in \{1, ..., N\}$, one for each subsystem Σ_I, and, that the structure of the diagnosis agents is exactly mirroring the decomposition (6.1), (6.2). The matrix $A_I \in \mathbb{R}^{n_I} \times \mathbb{R}^{n_I}$ and the function $f_I : \mathbb{R}^{p_I} \times \mathbb{R}^{\bar{p}_I} \times \mathbb{R}^{m_I} \mapsto \mathbb{R}^{n_I}$ are the known nominal function dynamics and the matrix $C_I \in \mathbb{R}^{p_I} \times \mathbb{R}^{n_I}$ is the known nominal output matrix of the I-th subsystem. The function f_I contains the known part of the interconnection function between the I-th and its neighboring subsystems. More specifically, the vectors $y_I^0(t)$, $\bar{y}_I^0(t)$ are defined as $y_I^0(t) \triangleq C_I x_I(t)$ and $\bar{y}_I^0(t) \triangleq \bar{C}_I \bar{x}_I(t)$, where $\bar{x}_I \in \mathbb{R}^{\bar{n}_I}$ and $\bar{C}_I \bar{x}_I \in \mathbb{R}^{\bar{p}_I}$ denote the state variables and the corresponding output variables, respectively, of the neighboring subsystems that affect the I-th subsystem. This indicates that f_I is a function of local variables $y_I^0(t)$ and, interconnection variables $\bar{y}_I^0(t)$ that are measurable as $y_I(t)$, $\bar{y}_I(t)$ respectively. The superscript ^{0} in $y_I^0(t)$, $\bar{y}_I^0(t)$ simply indicates the noiseless and sensor fault free measurements of $y_I(t)$, $\bar{y}_I(t)$ respectively. The vector function $\eta_I : \mathbb{R}^{n_I} \times \mathbb{R}^{\bar{n}_I} \times \mathbb{R}^{m_I} \times \mathbb{R}^+ \mapsto \mathbb{R}^{n_I}$ denotes the modeling uncertainty associated with the nominal dynamics and $\xi_I \in \mathbb{R}^{p_I}$ represents the measurement noise. The term $\beta_I^x(t - T_0^x)\phi_I(x, u_I)$ characterizes the process fault function dynamics affecting the I-th subsystem, including its time evolution. More specifically, the term $\phi_I : \mathbb{R}^n \times \mathbb{R}^{m_I} \mapsto \mathbb{R}^{n_I}$ represents the unknown fault function and the term $\beta_I^x(t - T_0^x) : \mathbb{R} \mapsto \mathbb{R}^+$ models the time evolution of the fault, where T_0^x is the unknown time of the fault occurrence. Note that the fault function ϕ_I may depend on the global state variable vector x and not only on the local state vector x_I allowing faults to be functions of the overall state vector and not only of the states that are available to the I-th subsystem. The term $\beta_I^y(t - T_0^y)\theta_I(t)$ characterizes the

Christodoulos Keliris

sensor fault, where $\beta_I^y : \mathbb{R} \mapsto \mathbb{R}^+$ models the time profile of the sensor fault which occurs at some unknown time T_0^y and, $\theta_I \in \mathbb{R}^{p_I}$ represents the unknown time-varying sensor fault. In this work, no particular modeling is considered for the time profile $\beta_I^x(t - T_0^x)$ of the process fault and $\beta_I^y(t - T_0^y)$ of the sensor fault. Generally, the time profiles can be used to model both abrupt and incipient faults. In this work we consider them to be zero prior to the respective fault occurrence and do not make any modeling considerations regarding the fault evolution after their occurrence, i.e. we only consider $\beta_I^x(t - T_0^x) = 0 \quad \forall t < T_0^x$ and $\beta_I^y(t - T_0^y) = 0 \quad \forall t < T_0^y$. In fact, the faults can be permanent, temporary or even intermittent.

In this work, subsystem Σ_J is said to to affect subsystem Σ_I (or in other words Σ_J is a neighbor of Σ_I), if the interconnection variables of Σ_I, i.e. $\bar{y}_I^0(t)$, contains at least one of the measurable output variables of Σ_J, i.e. $y_J^0(t)$.

The objective is to design and analyze a distributed fault detection approach where, a local fault detection agent $\mathcal{F}_I$ is associated with each subsystem Σ_I and receives local measurements u_I, y_I and partial information from neighboring fault detection agents $\mathcal{F}_J$. Each fault detection agent $\mathcal{F}_I$ is not connected to all other agents, but only to a subset of neighboring agents, thus constituting a distributed fault detection scheme. The analysis part does not aim only in the derivation of suitable detection thresholds but also, it aims in the investigation of the fault propagation issues. In this work, the notion of fault propagation does not mean the creation of new faults in interconnected subsystems as a result of a faulty behavior of a subsystem. Instead, it means the way a particular fault occurring in one subsystem affects neighboring interconnected subsystems (in other words, it is simply the propagation of the fault effects from a faulty subsystem to its interconnected subsystems it affects).

More specifically, the objective is to design a robust distributed fault detection scheme, for process and sensor faults, with enhanced fault detectability characteristics and, inherent fault isolation characteristics, that is able to provide information regarding the type and location of the fault that has occurred. The fault detectability enhancement is achieved through filtering by exploiting the noise suppression properties of filters in order to obtain tight detection thresholds. The enhanced fault isolation characteristics are integrated by the design of the scheme which utilizes the measurements instead of the state estimates. The purpose is to derive a high-level isolation scheme, which does not necessarily pinpoints the specific fault that has occurred, but rather infers some conclusions about the fault that has occurred by taking into consideration the decisions of the fault detection agents that monitor each subsystem. This information, which in some special cases could even lead to the identification

Christodoulos Kellinis

of a faulty sensor, can provide valuable information that can be used by a more advanced fault isolation scheme in order to greatly improve its performance by excluding potential fault scenarios. In the sequel, the terms fault isolation/diagnosis will be used according to the aforementioned basis.

Each fault detection agent contains an estimation model based on its subsystems' nominal dynamics that provides the state estimates and, utilizes filtering to derive the residual and threshold signals. Finally, each detection agent provides a binary decision regarding the detection of a fault in the subsystem it monitors. The decisions of all the fault detection agents are then exploited by the high-level isolation scheme to infer some information regarding the type/location of the fault that has occurred.

Generally, the distributed fault diagnosis scheme is composed by N local fault detection agents $\mathcal{F}_I$, $I = 1, \ldots, N$, one for each subsystem Σ_I. Each local fault detection agent $\mathcal{F}_I$ requires the input and output measurements of the subsystem Σ_I that is monitoring and also the measurements of all interconnecting subsystems Σ_J that are affecting Σ_I. Note that these last measurements are communicated by neighboring fault detection agents $\mathcal{F}_J$, and not by the subsystems Σ_J. Therefore, there is the need of communication between the fault detection agents depending on their interconnections which constitutes the scheme distributed. Note that, the information exchanged among the subsystems is constituted only by quantities ($\bar{y}_I^0(t)$) that are measurable with some uncertainty ($\bar{y}_I(t)$). Figure 6.1 illustrates the distributed fault diagnosis approach for the case of three subsystems Σ_1, Σ_2, Σ_3 where Σ_1 affects Σ_2 and Σ_2 affects Σ_3.

The fault diagnosis structure can also be considered as a hierarchical multi-agent diagnostic system composed of two layers: a lower and an upper layer. In the lower layer, a diagnostic agent $\mathcal{F}_I$ is associated to each subsystem Σ_I with the aim to detect process and sensor faults. Each fault detection agent $\mathcal{F}_I$ consists of a residual generator of the form (6.4), (6.5), (6.6) (to be given in the sequel), together with a fault detection decision logic relying on the comparison of the residual (6.6) to an adaptive threshold (6.9) (to be designed). In the upper layer, a fault isolation logic combines the decisions of the diagnostic agents, with the aim to infer further information regarding the type of the fault that has occurred (distinguish process and sensor faults) and its location.

In this thesis, we do not deal explicitly with the design of feedback controllers for selecting u_I. Instead, we consider the fault detection issue in the presence of process faults ϕ_I, sensor faults θ_I, modeling uncertainties η_I and measurement noise ξ_I. The proposed formulation allows for any controllers that achieve under healthy conditions some desired control

Christodoulos Kellis

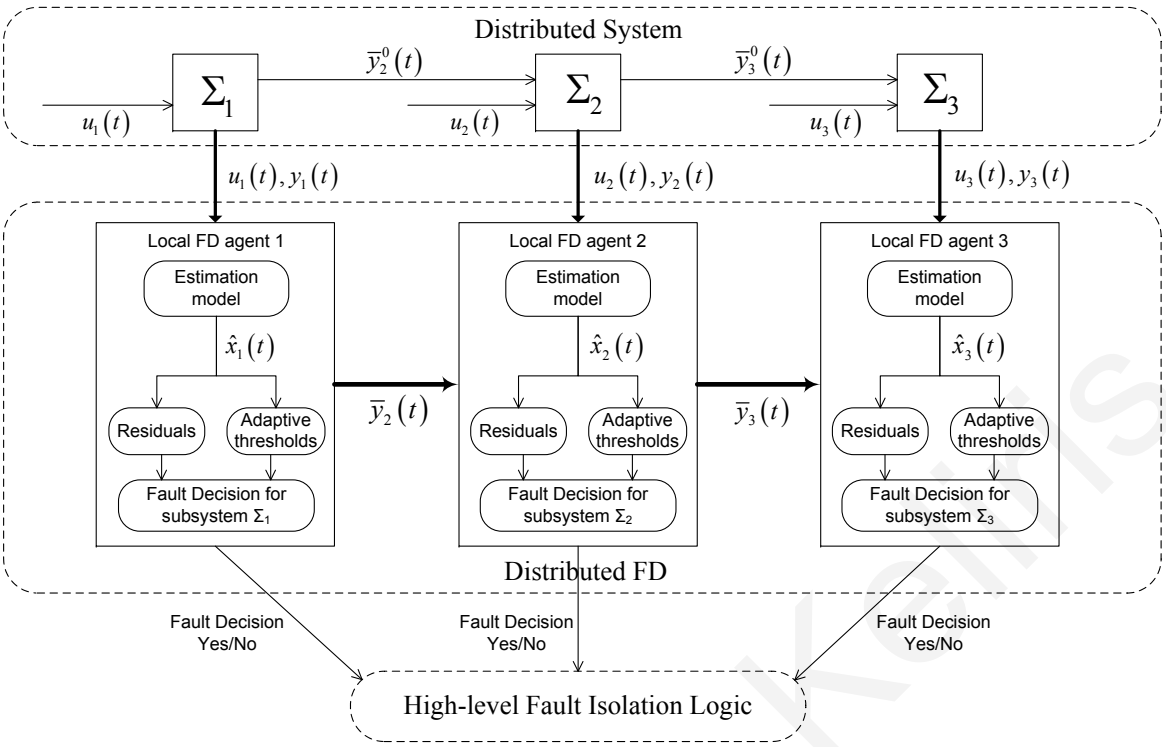


Figure 6.1: Distributed fault diagnosis approach for the case of three subsystems Σ_1, Σ_2, Σ_3 where Σ_1 affects Σ_2 and Σ_2 affects Σ_3.

objectives and does not depend on their structure. It is assumed that the controllers are able to retain the uniform boundedness of the state variables before and after the occurrence of a fault.

The following assumptions are used throughout the chapter:

***Assumption* 6.1.** *For each subsystem Σ_I, $I \in \{1, ..., N\}$, the local state variables $x_I(t)$ and the local inputs $u_I(t)$ belong to a known compact region $\mathcal{D}_{x_I}$ and $\mathcal{D}_{u_I}$ respectively before and after the occurrence of a fault, i.e. $x_I(t) \in \mathcal{D}_{x_I}$, $u_I(t) \in \mathcal{D}_{u_I}$ for all $t \geq 0$.*

***Assumption* 6.2.** *For each subsystem Σ_I, $I \in \{1, ..., N\}$ the pair (A_I, C_I) is detectable.*

***Assumption* 6.3.** *The modeling uncertainty η_I in each subsystem is an unstructured and unknown nonlinear function of x_I, $\bar{x}_I$, u_I and t but bounded by a known positive functional $\bar{\eta}_I$ under sensor fault-free operation, i.e.,*

$$\|\eta_I(x_I, \bar{x}_I, u_I, t)\| \leq \bar{\eta}_I(y_I, \bar{y}_I, u_I), \tag{6.3}$$

for all $t \in \mathbb{N}$ and for all $(x_I, \bar{x}_I, u_I) \in \mathcal{D}_I$, where $\bar{y}_I \in \mathbb{R}^{\bar{p}_I}$ is the noisy counterpart of $\bar{y}_I^0(t)$, i.e. $\bar{y}_I = \bar{y}_I^0(t) + \bar{\xi}_I$, $\bar{\xi}_I \in \mathbb{R}^{\bar{p}_I}$ and $\bar{\eta}_I(y_I, \bar{y}_I, u_I) \geq 0$ is a known bounding function in some compact region of interest $\mathcal{D}_I = \mathcal{D}_{x_I} \times \mathcal{D}_{\bar{x}_I} \times \mathcal{D}_{u_I} \subset \mathbb{R}^{n_I} \times \mathbb{R}^{\bar{n}_I} \times \mathbb{R}^{m_I}$.

Christodoulos Keliris

***Assumption* 6.4.** *The measurement noise belongs to a known compact region, i.e.* $\xi_I(t) \in \mathcal{D}_{\xi_I} \subset \mathbb{R}^{p_I}$, $\bar{\xi}_I \in \mathcal{D}_{\bar{\xi}_I} \subset \mathbb{R}^{\bar{p}_I}$.

Assumption 6.1 is required for well-posedness since in this work we do not address the control design and fault accommodation problem, but instead the fault detection problem. Assumption 6.2 is required for the design of a suitable observer to be used for the residual signal generation. Assumption 6.3 characterizes the class of modeling uncertainties being considered. The bound $\bar{\eta}_I$ is required in order to distinguish the effects between modeling uncertainty and faults. Assumption 6.4 is required in order to distinguish the effects between noise and faults.

6.2 Distributed Fault Detection

In this Section, the details of the proposed distributed scheme regarding the design of the residual and threshold signals are given, along with some practical considerations.

6.2.1 Residual Signal Generation

In this part, the residual signal generation in each fault detection agent is addressed by making use of filtering. For each subsystem Σ_I, we consider an estimation model, based on the known components of (6.1) under healthy mode of operation:

$$\hat{x}_I(t+1) = A_I\hat{x}_I(t) + f_I(y_I(t), \bar{y}_I(t), u_I(t)) + L_I(y_I(t) - \hat{y}_I(t)) \tag{6.4}$$

$$\hat{y}_I(t) = C_I\hat{x}_I(t), \tag{6.5}$$

where the gain matrix L_I is computed so that $(A_I - L_IC_I)$ is Schur stable, i.e. its eigenvalues lie in the open unit disc. Note that, since the pair (A_I, C_I) is detectable, according to Assumption 6.2, such L_I can always be determined. In order to simplify the presentation of the mathematical calculations, the initial condition $\hat{x}_I(0)$ is considered known as $\hat{x}_I(0) = x_I(0)$. In the case $x_I(0)$ is not exactly known, the discrepancy $x_I(0) - \hat{x}_I(0)$ will appear in the calculations which, as it will be shown later, is multiplied by exponentially decaying functions, and therefore it does not affect substantially the subsequent analysis.

In this work, the residual signal $r_I(t)$ to be used for fault detection in each subsystem Σ_I is given by

$$r_I(t) \triangleq H(z)\left[y_I(t) - \hat{y}_I(t)\right], \tag{6.6}$$

Christodoulos Keliris

where $H(z)$ is a p-th order, asymptotically stable filter with proper transfer function

$$H(z) = \frac{d_0 + d_1 z^{-1} + d_2 z^{-2} + \ldots + d_p z^{-p}}{1 + c_1 z^{-1} + \ldots + c_p z^{-p}}. \tag{6.7}$$

Note that the form of $H(z)$ allows both IIR and FIR types of digital filters. In addition, note that, for the residual generation, each measured variable $y_I^{(j)}$ (j-th component of y_I) is filtered by $H(z)$ in order to dampen the effect of measurement uncertainty $\xi_I(t)$, so that tighter detection thresholds can be obtained. Specifically, for each measured variable $y_I^{(j)}$, a corresponding residual $r_I^{(j)}$ and threshold $\bar{r}_I^{(j)}$ are generated and hence, all the measurements need to be filtered for the residual generation. On the other hand, the estimation model given by (6.4), (6.5) relies on the unfiltered measurements. In the proposed approach, the same filter $H(z)$ must be used within the fault detection agent $\mathcal{F}_I$ for filtering all the measurements y_I, and different filters $H(z)$ can be used by each fault detection agent. In this work, without loss of generality, we have considered that the same filter $H(z)$ is used by all detection agents.

The choice of a particular type of filter to be used is application dependent, and it is made according to the available a-priori knowledge on the noise properties. Usually, measurement noise is constituted by high frequency components and therefore the use of low-pass filter for dampening noise is well justified. On other occasions, one may want to focus the fault detectability on a prescribed frequency band of the measurement signals and hence, choose the filter accordingly. The particular selection criteria for choosing a suitable filter and its trade-offs are out of the scope of the present work and the reader is referred to the continuous-time case in Chapter 3 where a rigorous investigation of the filtering impact (according to the poles' location and filters' order) on the detection time is presented.

Since the filter $H(z)$ is asymptotically stable, for bounded measurement noise $\xi_I(t)$, the filtered measurement noise $\epsilon_{\xi_I}(t) \triangleq H(z)\,[\xi_I(t)]$ is bounded as follows:

$$\|\epsilon_{\xi_I}(t)\| \leq \bar{\epsilon}_{\xi_I}(t), \tag{6.8}$$

where $\bar{\epsilon}_{\xi_I}$ is a computable bounding function. Depending on the noise characteristics, $H(z)$ can be selected to reduce the bounding function $\bar{\epsilon}_{\xi_I}$. It is important to note that filtering is primarily used to mitigate the effect of measurement noise and aid in the derivation of tighter thresholds, thus enhancing fault detectability (see [49]).

Remark 6.1. *It is important to note that in the nonlinear function f_I, the measurements y_I and $\bar{y}_I$ are used instead of the estimates $C_I\hat{x}_I$ and $\bar{C}_I\hat{\bar{x}}_I$. This is crucial for the derivation of the high-level fault isolation scheme, since in the case of a process fault its effects are*

Christodoulos Kallis

contained in the measurements, and as it will be shown, the fault can only be detected by the agent monitoring the subsystem that the fault has occurred. On the other hand, when a sensor fault occurs, it can affect neighboring detection agents through the communicated measurements of the interconnection variables, i.e. $\bar{y}_I$ that contain the sensor fault and hence the fault can also be detected by interconnected detection agents. In [98], which investigated a decentralized fault detection scheme for process faults, the estimation model used the estimates in the interconnection functions among subsystems (instead of the measurements as in this work) in order to enhance fault detectability by allowing the interconnected agents to also be able to detect the fault. In this work, the measurements are used instead of the estimates, in order to enhance fault isolability.

6.2.2 Adaptive Detection Threshold

In this part, suitable detection thresholds that guarantee no false alarms are derived. Filtering is also integrated in the design in order to attenuate the measurement noise effects and aid in the derivation of tighter detection thresholds. The filtering use is treated as a linear state transformation, which essentially converts the system and estimator dynamics to their filtered counterparts, allowing easier mathematical manipulation of the expressions.

In the following, we will use the filtered state variable $x_{I,f}(t) \triangleq H(z)\,[x_I(t)]$, and the filtered state estimate $\hat{x}_{I,f}(t) \triangleq H(z)\,[\hat{x}_I(t)]$. The threshold design is based on the derivation of a suitable bound on the filtered state estimation error $x_{I,f}(t) - \hat{x}_{I,f}(t)$. Also, let $h(t)$ be the impulse response associated with $H(z)$; i.e. $h(t) \triangleq \mathcal{Z}^{-1}\,[H(z)]$, so that $x_{I,f}(t)$ can be written as

$$x_{I,f}(t) = \sum_{k=0}^{t} h(k) x_I(t-k).$$

In this work, the detection decision of a fault in the overall system is made when $|r_I^{(j)}(t)| > \bar{r}_I^{(j)}(t)$ at some time t for at least one component $j \in \{1, 2, \ldots, p_I\}$ in any local subsystem Σ_I, where $\bar{r}_I^{(j)}(t)$ is the detection threshold given by

$$\bar{r}_I^{(j)}(t) \triangleq \sum_{k=0}^{t-1} \alpha_{I,j} \delta_{I,j}^{t-1-k} \bar{\chi}_I(k) + \bar{\epsilon}_{\xi_I}(t), \tag{6.9}$$

where

$$\bar{\chi}_I(t) \triangleq \bar{H}(z)\left[\bar{\eta}_I(y_I(t), \bar{y}_I(t), u_I(t)) + \bar{\Delta} f_I\right] + \|L_I\| \bar{\epsilon}_{\xi_I}(t), \tag{6.10}$$

Christodoulos Keliris

$\bar{H}(z)$ is a filter with impulse response $\bar{h}(t) \geq |h(t)|$ for all $t \geq 0$,

$$\bar{\Delta} f_I \triangleq \sup_{\substack{(x_I, \bar{x}_I, u_I) \in \mathcal{D}_I \\ (\xi_I, \bar{\xi}_I) \in \mathcal{D}_{\xi_I} \times \mathcal{D}_{\bar{\xi}_I}}} \| f_I\big(C_I x_I, \bar{C}_I \bar{x}_I, u_I\big) - f_I\big(C_I x_I + \xi_I, \bar{C}_I \bar{x}_I + \bar{\xi}_I, u_I\big) \|, \tag{6.11}$$

and where the constants $\alpha_{I,j} > 0$ and $0 < \delta_{I,j} < 1$ are selected so that the following inequality holds:

$$\|C_I^{(j)} A_{I,0}^t\| \leq \alpha_{I,j} \delta_{I,j}^t \leq \|C_I^{(j)}\| \|A_{I,0}\|^t, \tag{6.12}$$

with $A_{I,0} \triangleq A_I - L_I C_I$. Finally, $C_I^{(j)}$ denotes the j-th row of the matrix C_I. Note that, since $A_{I,0}$ is Schur stable, suitable constants $\alpha_{I,j}$, $\delta_{I,j}$ do exist [27].

Note that, the threshold (6.9) can be implemented using linear filtering techniques as:

$$\bar{r}_I^{(j)}(t) = \frac{\alpha_{I,j}}{1 - \delta_{I,j} z^{-1}} \left[\bar{\chi}_I(t-1)\right] + \bar{\epsilon}_{\xi_I}(t). \tag{6.13}$$

Figure 6.2 illustrates the implementation of the fault detection scheme for the detection agent $\mathcal{F}_I$ resulting from equations (6.7), (6.4), (6.5), (6.6) and (6.13).

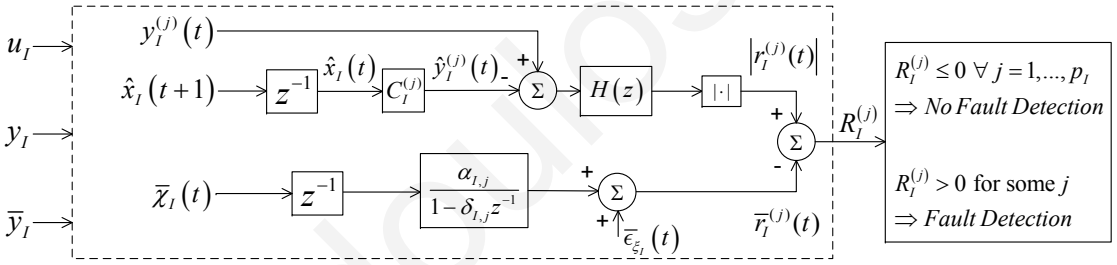


Figure 6.2: Fault detection scheme of the detection agent $\mathcal{F}_I$.

In the absence of any faults, the residual signal $r_I^{(j)}(t)$ given in (6.6) is always bounded by the detection threshold $\bar{r}_I^{(j)}(t)$ given by (6.9). The fault detection concept is formalized in the following Lemma.

Lemma 6.1. *Consider a distributed system made of N subsystems Σ_I given by* (6.1)*,* (6.2)*. In the absence of any faults, the residuals $r_I^{(j)}(t)$ given by* (6.6)*, where $\hat{y}_I^{(j)}$ are given by* (6.4) *and* (6.5)*, are bounded by the detection thresholds $\bar{r}_I^{(j)}(t)$, given by* (6.9)*, thus guaranteeing that no false alarms are issued by the fault detection scheme.*

Proof. The filtered state estimate $x_{I,f}(t+1)$ can be written as:

$$\begin{aligned} x_{I,f}(t+1) &= \sum_{k=0}^{t+1} h(t+1-k) x_I(k) \\ &= \sum_{k=1}^{t+1} h(t+1-k) x_I(k) + h(t+1) x_I(0), \end{aligned}$$

Christodoulos Keliris

and, by using the change of variables $k = i + 1$, it becomes

$$x_{I,f}(t+1) = \sum_{i=0}^{t} h(t-i)x_I(i+1) + h(t+1)x_I(0). \tag{6.14}$$

i) Prior to the possible occurrence of a process fault, by using (6.1), (6.14) can be written as:

$$\begin{aligned} x_{I,f}(t+1) &= \sum_{i=0}^{t} h(t-i)\big[A_I x_I(i) + f_I(y_I^0(i), \bar{y}_I^0(i), u_I(i)) + \eta_I(x_I(i), \bar{x}_I(i), u_I(i), i)\big] \\ &\quad + h(t+1)x_I(0) \\ &= A_I x_{I,f}(t) + H(z)\left[f_I(y_I^0(t), \bar{y}_I^0(t), u_I(t))\right] + H(z)\left[\eta_I(x_I(t), \bar{x}_I(t), u_I(t), t)\right] \\ &\quad + h(t+1)x_I(0). \end{aligned} \tag{6.15}$$

Similarly to the derivation of (6.15), the filtered state estimate dynamics $\hat{x}_{I,f}(t+1)$ by using (6.4) satisfies:

$$\begin{aligned} \hat{x}_{I,f}(t+1) &= A_I \hat{x}_{I,f}(t) + H(z)\big[f_I\big(y_I(t), \bar{y}_I(t), u_I(t)\big)\big] + L_I\big(y_{I,f}(t) - \hat{y}_{I,f}(t)\big) \\ &\quad + h(t+1)\hat{x}_I(0), \end{aligned} \tag{6.16}$$

where $y_{I,f}(t) \triangleq H(z)[y_I(t)]$ and $\hat{y}_{I,f}(t) \triangleq H(z)[\hat{y}_I(t)] = C_I \hat{x}_{I,f}(t)$.

ii) Prior to the possible occurrence of a sensor fault, $y_{I,f}(t) = C_I x_{I,f}(t) + \epsilon_{\xi_I}(t)$ and, by using (6.15), (6.16) and some algebra, the filtered state estimation error $\tilde{x}_{I,f}(t) \triangleq x_{I,f}(t) - \hat{x}_{I,f}(t)$ satisfies

$$\tilde{x}_{I,f}(t+1) = A_{I,0}\tilde{x}_{I,f}(t) + \chi_I(t), \tag{6.17}$$

where

$$\chi_I(t) \triangleq H(z)\left[\eta_I(x_I(t), \bar{x}_I(t), u_I(t), t) + \Delta f_I(t)\right] - L_I \epsilon_{\xi_I}(t) + h(t+1)(x_I(0) - \hat{x}_I(0)), \tag{6.18}$$

$$\Delta f_I(t) \triangleq f_I\big(y_I^0(t), \bar{y}_I^0(t), u_I(t)\big) - f_I\big(y_I^0(t) + \xi_I(t), \bar{y}_I^0(t) + \bar{\xi}_I(t), u_I(t)\big). \tag{6.19}$$

The solution of (6.17) is

$$\tilde{x}_{I,f}(t) = A_{I,0}^t \tilde{x}_{I,f}(0) + \sum_{k=0}^{t-1} A_{I,0}^{t-1-k} \chi_I(k). \tag{6.20}$$

Note that in (6.18), $h(t+1)(x_I(0) - \hat{x}_I(0)) = 0$ because of $x_I(0) = \hat{x}_I(0)$, and note that the term $\tilde{x}_{I,f}(0)$ in (6.20) is also zero since $\tilde{x}_{I,f}(0) = x_{I,f}(0) - \hat{x}_{I,f}(0) = h(0)(x_I(0) - \hat{x}_I(0)) = 0$. In the case $\hat{x}_I(0) \neq x_I(0)$, the aforementioned terms decay exponentially to zero and hence do not affect substantially the subsequent analysis. More specifically, the

Christodoulos Keliris

term $A_{I,0}^t \tilde{x}_{I,f}(0)$ in (6.20) decays exponentially to zero because $A_{I,0}$ is Schur stable, and the term $h(t+1)(x_I(0) - \hat{x}_I(0))$ in (6.18) decays also exponentially to zero because the impulse response $h(t)$ of an asymptotically stable filter is exponentially decaying (in the case of IIR filter, the FIR filter case is trivial).

Now, by using (6.2) and (6.5) the residual (6.6) prior to any fault satisfies $r_I(t) = C_I \tilde{x}_{I,f}(t) + \epsilon_{\xi_I}(t)$ and, by using (6.20) and $\tilde{x}_{I,f}(0) = 0$, it becomes

$$r_I(t) = \sum_{k=0}^{t-1} C_I A_{I,0}^{t-1-k} \chi_I(k) + \epsilon_{\xi_I}(t). \tag{6.21}$$

By taking the absolute value component-wise and using the triangle inequality, the j-th element of $r_I(t)$, i.e. $r_I^{(j)}(t)$ satisfies

$$\begin{aligned} |r_I^{(j)}(t)| &\leq \left| \sum_{k=0}^{t-1} C_I^{(j)} A_{I,0}^{t-1-k} \chi_I(k) \right| + |\epsilon_{\xi_I}^{(j)}(t)| \\ &\leq \sum_{k=0}^{t-1} \|C_I^{(j)} A_{I,0}^{t-1-k}\| \|\chi_I(k)\| + |\epsilon_{\xi_I}^{(j)}(t)|. \end{aligned} \tag{6.22}$$

Moreover, using (6.12) and the fact that $|\epsilon_{\xi_I}^{(j)}(t)| \leq \|\epsilon_{\xi_I}(t)\| \leq \bar{\epsilon}_{\xi_I}(t)$, (6.22) becomes

$$|r_I^{(j)}(t)| \leq \sum_{k=0}^{t-1} \alpha_{I,j} \delta_{I,j}^{t-1-k} \|\chi_I(k)\| + \bar{\epsilon}_{\xi_I}(t). \tag{6.23}$$

Now, consider the term $\chi_I(t)$ which satisfies

$$\begin{aligned} \|\chi_I(t)\| &= \|H(z)\left[\eta_I(x_I, \bar{x}_I, u_I, t) + \Delta f_I(t)\right] - L_I \epsilon_{\xi_I}(t)\| \\ &\leq \|H(z)\left[\eta_I(x_I, \bar{x}_I, u_I, t) + \Delta f_I(t)\right]\| + \|L_I \epsilon_{\xi_I}(t)\| \\ &\leq \sum_{k=0}^{t} |h(t-k)| \|\eta_I(x_I(k), \bar{x}_I(k), u_I(k), k)\| + \sum_{k=0}^{t} |h(t-k)| \|\Delta f_I(k)\| \\ &\quad + \|L_I\| \|\epsilon_{\xi_I}(t)\| \\ &\leq \bar{\chi}_I(t) \end{aligned} \tag{6.24}$$

where $\bar{\chi}_I(t)$ is the bounding function given by (6.10).

Finally, by using (6.23), (6.24) and the bound $\bar{\chi}_I(t)$, we obtain $|r_I^{(j)}(t)| \leq \bar{r}_I^{(j)}(t)$, where the detection threshold $\bar{r}_I^{(j)}(t)$ is given by (6.9), thus concluding the proof.

□

From a practical viewpoint, the implementation of the threshold $\bar{r}_I$ requires the bound $\bar{\Delta} f_I$ given in (6.11). One approach to derive this bound is to consider a local Lipschitz condition, i.e.:

$$\|f_I\big(C_I x_I, \bar{C}_I \bar{x}_I, u_I\big) - f_I\big(C_I x_I + \xi_I, \bar{C}_I \bar{x}_I + \bar{\xi}_I, u_I\big)\| \leq L_{f_I} \|[\xi_I \quad \bar{\xi}_I]^\top\|,$$

Christodoulos Kellris

where L_{f_I} is the Lipschitz constant for the function f_I with respect to $(x_I, \bar{x}_I)$ in the region $\mathcal{D}_{x_I} \times \mathcal{D}_{\bar{x}_I}$. Therefore, by using a uniform bound on the measurement noise (see Assumption 6.4), i.e. $\|\xi_I(t)\| \leq \xi_{I,b}$, $\|\bar{\xi}_I(t)\| \leq \bar{\xi}_{I,b} \quad \forall t \in \mathbb{N}$, then we can derive the bound $\bar{\Delta} f_I$.

Filtering is primarily used to dampen the measurement noise and allow the derivation of tighter detection thresholds. Filtering can also be proved beneficial for dampening the mismatch function $\Delta f_I(t)$ which results due to the measurement noise and therefore further enhance fault detectability. Among the various filters $H(z)$ one can select, some may lead to less conservative detection thresholds. The derivation of potentially tighter thresholds is obtained by writing the total uncertainty term $\chi_I(t)$ given by (6.18) as

$$\chi_I(t) = H(z)\left[\eta_I(x_I, \bar{x}_I, u_I, t)\right] + \epsilon_{\Delta f_I}(t) - L_I \epsilon_{\xi_I}(t), \tag{6.25}$$

where $\epsilon_{\Delta f_I}(t) \triangleq H(z)\left[\Delta f_I(t)\right]$, and by making the following assumption.

***Assumption* 6.5.** *In the absence of a sensor fault, the filtered function mismatch term $\epsilon_{\Delta f_I}(t)$ is bounded by a computable positive function $\bar{\epsilon}_{\Delta f_I}(t)$; i.e., for all $t \in \mathbb{N}$,*

$$\|\epsilon_{\Delta f_I}(t)\| \leq \bar{\epsilon}_{\Delta f_I}(t). \tag{6.26}$$

Assumption 6.5 is based on the fact that filtering dampens the error effect of measurement noise present in the function mismatch term $\Delta f_I(t)$. A suitable selection of $\bar{\epsilon}_{\Delta f_I}$ can be made through the use of simulations (i.e. Monte Carlo methods) by filtering the function mismatch term $\Delta f_I(t)$ using the known nominal function dynamics and the available noise characteristics (recall that the measurement noise is assumed to take values in a compact set, see Assumption 6.4).

In this case, the detection threshold $\bar{r}_I^{(j)}(t)$ is still given by (6.9), but $\bar{\chi}_I(t)$ is now given by:

$$\bar{\chi}_I(t) = \quad \bar{H}(z)\left[\bar{\eta}_I(y_I(t), \bar{y}_I(t), u_I(t))\right] + \bar{\epsilon}_{\Delta f_I}(t) + \|L_I\| \bar{\epsilon}_{\xi_I}(t). \tag{6.27}$$

As a result, the detection threshold by using (6.27) can be less conservative than by using (6.10). More information regarding this can be found in [49].

6.2.3 Selection of filter $\bar{H}(z)$

In this part, we give more details on the selection of a suitable filter $\bar{H}(z)$ which is required for the implementation of the detection threshold. As stated before, its impulse response must satisfy $|h(t)| \leq \bar{h}(t)$ for all $t \geq 0$. In the case where the impulse response $h(t)$ is

Christodoulos Kellinis

non-negative, the selection $\bar{H}(z) = H(z)$ is trivial. Sufficient conditions for non-negative impulse response for a class of discrete-time transfer functions are given in [59]. In the following, we briefly illustrate two simple methods for choosing $\bar{H}(z)$, one considering $H(z)$ as a digital IIR filter and the other one as a FIR filter.

First we consider the case where $H(z)$ is an IIR filter. As stated earlier, the impulse response $h(t)$ of a proper and asymptotically stable transfer function $H(z)$ converges to zero exponentially fast. Therefore, there exist $\kappa > 0$, $\lambda \in [0, 1)$ such that for all $t \in \mathbb{N}$ the following inequality holds: $|h(t)| \leq \kappa\lambda^t$. Since $|h(t)| \leq \bar{h}(t)$ must hold, the impulse response $\bar{h}(t)$ can be selected as $\bar{h}(t) = \kappa\lambda^t$ and thus $\bar{H}(z) = \frac{\kappa}{1-\lambda z^{-1}}$.

Now, let's consider the case in which $H(z)$ is a FIR filter. Let $H(z)$ be a p-th order FIR filter given by $H(z) = \sum_{k=0}^{p} d_k z^{-k}$. Therefore, $\bar{h}(t)$ can be selected as $\bar{h}(t) = |h(t)|$ which leads to the FIR filter $\bar{H}(z) = \sum_{k=0}^{p} |d_k| z^{-k}$.

Generally, in fault detection schemes it is difficult to find a suitable good balance in selecting the threshold. If the threshold is too high (conservative threshold) then some faults may go undetected. If the threshold is too low, then this may result in false alarms. The following Remark discusses the sources of possible conservativeness in the proposed scheme.

Remark 6.2. *The designed fault detection scheme guarantees that in the absence of a fault, the residual $r_I(t)$ given by* (6.6) *is uniformly bounded by the detection threshold $\bar{r}_I(t)$ given by* (6.9)*, which requires the bound $\bar{\chi}_I(t)$. The main sources of conservativeness in the designed threshold are included in the overall bound $\bar{\chi}_I(t)$, which can be broken up into the following four components: a) the filter $\bar{H}(z)$, b) the bound on the modeling uncertainty $\eta_I(x_I, \bar{x}_I, u_I)$, c) the bound on the filtered noise $\epsilon_{\xi_I}(t)$ and d) the bound on the function mismatch term $\Delta f_I(t)$. At first, the filter $\bar{H}(z)$, which is required to satisfy $|h(t)| \leq \bar{h}(t)$ for all $t \geq 0$, may impose some conservativeness but, this can be avoided by selecting the filter $H(z)$ used for filtering the measurements to have a non-negative impulse response so that the same filter can be used for the threshold derivation, i.e. $\bar{H}(z) = H(z)$. In this case, no conservativeness in the threshold is added. The second source of conservativeness stems from the bound on $\eta_I(x_I, \bar{x}_I, u_I)$, which according to Assumption 6.3, is bounded by a known function $\bar{\eta}_I(y_I, \bar{y}_I, u_I)$. This is required in order to distinguish the effects from modeling uncertainty and faults so that no false alarms are introduced. In practice, the system can be modeled more accurately in certain regions of the state space and therefore, the fact that the bound $\bar{\eta}_I$ is a function of y_I, $\bar{y}_I$ and u_I provides more flexibility by allowing the designer to take into consideration any prior knowledge of the system. Regarding the conservativeness*

Christodoulos Kallis

imposed by the use of the bound on the filtered noise $\epsilon_{\xi_I}(t)$, please note that the bound $\bar{\epsilon}_{\xi_I}(t)$ is multiplied with $\|L_I\|$ in (6.10) *and hence, the conservativeness is significantly reduced in comparison to the case in which no filtering is used (in that case, the bound $\bar{\xi}_I(t)$ on the noise (i.e. $\|\xi_I(t)\| \leq \bar{\xi}_I(t)$) would be multiplied with $\|L_I\|$, leading to more conservative thresholds). Finally, the last source of conservativeness in the threshold is introduced by the bound on the mismatch function Δf_I given by* (6.19) *for which the bound $\bar{\Delta} f_I$ given by* (6.11) *is required. As discussed earlier, one way to derive this bound is through the use of the Lipschitz assumption. The use of filtering can aid in the elimination of (some of) the conservativeness imposed by the bound $\bar{\Delta} f_I$, by exploiting the noise suppression properties through Assumption 6.5 and by using the bound $\bar{\chi}_I(t)$ given by* (6.27).

In Section 6.3, fault detectability conditions for the aforementioned fault detection scheme are addressed. The conditions given in Section 6.3 refer to the case of a fault occurring in subsystem Σ_I and being detected by its respective local fault detection agent $\mathcal{F}_I$. In Section 6.4, fault propagation from one subsystem to the other is investigated by examining the way the fault effects appear and affect neighboring interconnected subsystems.

6.3 Local fault detectability analysis

The fault detectability analysis constitutes a theoretical result that characterizes quantitatively (and in implicit way) the class of faults detectable by the proposed scheme. In order to derive the fault detectability conditions, we take into consideration the distinct occurrence of a process fault ϕ_I at an unknown time $t = T_0^x$ or the occurrence of a sensor fault θ_I at an unknown time $t = T_0^y$.

Theorem 6.1 (Local Process Fault Detectability)**.** *Consider the nonlinear interconnected system (6.1), (6.2) with the distributed fault detection scheme described in (6.4), (6.5), (6.6), (6.9) and (6.10) . A process fault in the I-th subsystem occurring at $t = T_0^x$ is detectable by the respective local fault detection agent $\mathcal{F}_I$ if the filtered process fault function $\phi_{I,f}\big(x(t), u_I(t), t\big) \triangleq H(z)[\beta_I^x(t - T_0^x)\phi_I(x(t), u_I(t))]$ satisfies the following inequality at some time $t > T_0^x$, for some $j = 1, 2, \ldots, p_I$:*

$$\left| \sum_{k=T_0^x}^{t-1} C_I^{(j)} A_{I,0}^{t-1-k} \phi_{I,f}\big(x(k), u_I(k), k\big) \right| > 2\bar{r}_I^{(j)}(t). \tag{6.28}$$

Proof. In the presence of a process fault that occurs in the I-th subsystem at $t = T_0^x$, (6.20)

Christodoulos Keliris

becomes:

$$\tilde{x}_{I,f}(t) = A_{I,0}^{t}\tilde{x}_{I,f}(0) + \sum_{k=0}^{t-1} A_{I,0}^{t-1-k}\big(\chi_I(k) + \phi_{I,f}\big(x(k), u_I(k), k\big)\big),$$

and by using $\tilde{x}_{I,f}(0) = 0$, the residual $r_I(t) = C_I\tilde{x}_{I,f}(t) + \epsilon_{\xi_I}(t)$ becomes (similarly to (6.21))

$$r_I(t) = \sum_{k=0}^{t-1} C_I A_{I,0}^{t-1-k}\big(\chi_I(k) + \phi_{I,f}\big(x(k), u_I(k), k\big)\big) + \epsilon_{\xi_I}(t).$$

By using the triangle inequality, the j-th element of $r_I(t)$ for $t > T_0^x$ satisfies:

$$|r_I^{(j)}(t)| \geq -\left|\sum_{k=0}^{t-1} C_I^{(j)} A_{I,0}^{t-1-k}\chi_I(k)\right| - |\epsilon_{\xi_I}^{(j)}(t)| + \left|\sum_{k=T_0^x}^{t-1} C_I^{(j)} A_{I,0}^{t-1-k}\phi_{I,f}\big(x(k), u_I(k), k\big)\right|. \tag{6.29}$$

Following a similar procedure as in the derivation of the detection threshold (6.9), (6.29) becomes

$$|r_I^{(j)}(t)| \geq -\bar{r}_I^{(j)}(t) + \left|\sum_{k=T_0^x}^{t-1} C_I^{(j)} A_{I,0}^{t-1-k}\phi_{I,f}\big(x(k), u_I(k), k\big)\right|.$$

For fault detection, the inequality $|r_I^{(j)}(t)| > \bar{r}_I^{(j)}(t)$ must hold for some $j = 1, 2, \ldots, p_I$, so the final fault detectability condition given in (6.28) is obtained. □

Theorem 6.2 (Local Sensor Fault Detectability)**.** *Consider the nonlinear interconnected system (6.1), (6.2) with the distributed fault detection scheme described in (6.4), (6.5), (6.6), (6.10) and (6.9). A sensor fault in the I-th subsystem occurring at $t = T_0^y$ is detectable by the respective local fault detection agent $\mathcal{F}_I$ if the filtered sensor fault function*

$$\theta_{I,f}(t) \triangleq H(z)[\beta_I^y(t - T_0^y)\theta_I(t)],$$

and the following mismatch function $\Delta f_I'$ due to the sensor fault

$$\Delta f_I'(t) \triangleq f_I(y_I^0(t), \bar{y}_I^0(t), u_I(t)) - f_I(y_I^0(t) + \xi_I(t) + \beta_I^y(t - T_0^y)\theta_I(t), \bar{y}_I^0(t) + \bar{\xi}_I(t), u_I(t)),$$

satisfy the following inequality at some time $t > T_0^y$, for some $j = 1, 2, \ldots, p_I$:

$$\left|\theta_{I,f}^{(j)}(t) + \sum_{k=0}^{t-1} C_I^{(j)} A_{I,0}^{t-1-k}\big(H(z)[\Delta g_I'(k)] - L_I\theta_{I,f}(k)\big)\right| > \bar{r}_I^{(j)}(t) + \bar{\epsilon}_{\xi_I}(t) + \sum_{k=0}^{t-1} \alpha_{I,j}\delta_{I,j}^{t-1-k}\big(\|H(z)[\eta_I(x_I, \bar{x}_I, u_I, k)]\| + \|L_I\|\bar{\epsilon}_{\xi_I}(k)\big). \tag{6.30}$$

Christodoulos Keliris

Proof. In the presence of a sensor fault that occurs in the I-th subsystem at $t = T_0^y$, (6.20) becomes

$$\tilde{x}_{I,f}(t) = A_{I,0}^{t}\tilde{x}_{I,f}(0) + \sum_{k=0}^{t-1} A_{I,0}^{t-1-k}\chi_I'(k), \tag{6.31}$$

where

$$\chi_I'(t) \triangleq H(z)\big[\eta_I(x_I(t), \bar{x}_I(t), u_I(t), t) + \Delta f_I'(t)\big] - L_I\epsilon_{\xi_I}(t) - L_I\theta_{I,f}(t).$$

Let's define $v_{I,f}(t) \triangleq H(z)[\Delta g_I'(t)] - L_I\theta_{I,f}(t)$. After the occurrence of a sensor fault, the residual (6.6) becomes $r_I(t) = C_I\tilde{x}_{I,f}(t) + \epsilon_{\xi_I}(t) + \theta_{I,f}(t)$ and by using (6.31) with $\tilde{x}_{I,f}(0) = 0$, the residual is written as:

$$\begin{aligned} r_I(t) &= \sum_{k=0}^{t-1} C_I A_{I,0}^{t-1-k}\chi_I'(k) + \epsilon_{\xi_I}(t) + \theta_{I,f}(t) \\ &= \sum_{k=0}^{t-1} C_I A_{I,0}^{t-1-k}\big(H(z)[\eta_I(x_I(k), \bar{x}_I(k), u_I(k), k)] - L_I\epsilon_{\xi_I}(k)\big) \\ &\quad + \sum_{k=0}^{t-1} C_I A_{I,0}^{t-1-k} v_{I,f}(k) + \epsilon_{\xi_I}(t) + \theta_{I,f}(t) \end{aligned}$$

By using the triangle inequality, the j-th element of $r_I(t)$ for $t > T_0^y$ satisfies:

$$\begin{aligned} |r_I^{(j)}(t)| &\geq \left|\theta_{I,f}^{(j)}(t) + \sum_{k=0}^{t-1} C_I^{(j)} A_{I,0}^{t-1-k} v_{I,f}(k)\right| - |\epsilon_{\xi_I}^{(j)}(t)| \\ &\quad - \left|\sum_{k=0}^{t-1} C_I^{(j)} A_{I,0}^{t-1-k}\big(H(z)[\eta_I(x_I(k), \bar{x}_I(k), u_I(k), k)] - L_I\epsilon_{\xi_I}(k)\big)\right| \\ &\geq \left|\theta_{I,f}^{(j)}(t) + \sum_{k=0}^{t-1} C_I^{(j)} A_{I,0}^{t-1-k} v_{I,f}(k)\right| - \bar{\epsilon}_{\xi_I}(t) \\ &\quad - \sum_{k=0}^{t-1} \alpha_{I,j}\delta_{I,j}^{t-1-k}\big(\|H(z)[\eta_I(x_I(k), \bar{x}_I(k), u_I(k), k)]\| + \|L_I\|\bar{\epsilon}_{\xi_I}(k)\big), \end{aligned}$$

For fault detection, the inequality $|r_I^{(j)}(t)| > \bar{r}_I^{(j)}(t)$ must hold for some $j = 1, 2, \ldots, p_I$, so the final fault detectability condition given in (6.30) is obtained. □

Theorems 6.1 and 6.2 provide sufficient conditions for the implicit characterization of certain classes of faults that can be detected by the proposed fault detection scheme. Clearly, the fault functions $\phi_I(x, u_I)$ and θ_I are typically unknown and therefore these conditions cannot be checked a-priori.

Remark 6.3. *The use of filtering is of crucial importance in order to derive tight detection thresholds that guarantee no false alarms (see [49]). As it can be seen in the detectability*

Christoforos Kellis

conditions given by (6.28)*,* (6.30) *the detection of the fault depends on the filtered process fault function ϕ_I and filtered sensor fault θ_I and as a result, the selection of the filter is very important. Therefore, some filter selections may lead to less conservative thresholds than others.*

6.4 Fault propagation

In this section, fault propagation from one subsystem to the other is investigated and further intuition regarding the isolation properties of the proposed fault diagnosis scheme is obtained. The notion of fault propagation does not mean the creation of additional faults to neighboring subsystems as a result of a faulty behavior in one of them. Instead, it means the way the fault effects in one subsystem appear and affect its neighboring subsystems. More specifically, we consider a fault that occurs in subsystem Σ_J which affects Σ_I and investigate the possibility of fault detection, not by the local fault detection agent $\mathcal{F}_J$ (which may obviously detect the fault), but by the agent $\mathcal{F}_I$.

The following lemma summarizes the main findings of this analysis.

Lemma 6.2. *Consider a distributed system made of N subsystems Σ_I given by* (6.1)*,* (6.2)*. The distributed fault detection scheme described by the estimation model* (6.4)*,* (6.5)*, the residual signals $r_I(t)$ given by* (6.6) *and the detection thresholds $\bar{r}_I(t)$ given by* (6.9) *guarantees that:*

(a) a process fault occurring in subsystem Σ_J which affects Σ_I can only be detected by its corresponding fault detection agent $\mathcal{F}_J$ and not by the detection agent $\mathcal{F}_I$.

(b) a sensor fault occurring in Σ_J which affects Σ_I can be detected by either the corresponding detection agent $\mathcal{F}_J$ or the detection agent $\mathcal{F}_I$.

Proof. The case of the fault (process or sensor) occurring in subsystem Σ_J and being detectable by its corresponding detection agent $\mathcal{F}_J$ has been investigated in Section 6.2 and respective detectability conditions were given in Section 6.3. In the sequel, we investigate the possibility of detection of a fault that occurs in Σ_J which affects Σ_I by the detection agent $\mathcal{F}_I$.

(a) At first, lets consider fault propagation in the case of a process fault. In this case, the process fault effects in Σ_J are propagated to Σ_I through the interconnection variables $\bar{C}_I\bar{x}_I$ (see (6.1)) and to $\mathcal{F}_I$ through the measurements $\bar{y}_I$ (communicated by $\mathcal{F}_J$, see (6.4)). For

Christodoulos Kallis

easier visual indication of the process fault effects that are contained in the interconnection variables $\bar{x}_I$ and the measurements of the output interconnection variables $\bar{y}_I$, we denote them as $\bar{x}_{I,P}$ and $\bar{y}_{I,P}$ respectively. Note that $\bar{y}_{I,P}(t) = \bar{C}_I \bar{x}_{I,P}(t) + \bar{\xi}_I(t)$. In the case of a process fault occurring in Σ_J which affects Σ_I, the dynamics of Σ_I are given by (6.1), (6.2) and the estimation model of the local fault detection agent $\mathcal{F}_I$ by (6.4), (6.5). In the aforementioned equations $\bar{x}_I$ and $\bar{y}_I$ are now indicated by $\bar{x}_{I,P}$ and $\bar{y}_{I,P}$ respectively. Note that, both the dynamics of Σ_I and the estimation model of the local fault detection agent $\mathcal{F}_I$ are affected by the process fault effects that occurred in Σ_J. Following the same analysis as in the proof of Lemma 6.1, the filtered estimation error still satisfies (6.17)-(6.19) (the fault effects enter implicitly through the interconnection variables and their measurements), which are rewritten below with explicit indication of the fault effect through $\bar{x}_{I,P}$ ($\bar{y}_{I,P}$ is expressed in terms of $\bar{x}_{I,P}$):

$$
\begin{aligned}
&\tilde{x}_{I,f}(t+1) = A_{I,0}\tilde{x}_{I,f}(t) + \chi_{I,P}(t) \\
&\chi_{I,P}(t) \triangleq H(z)\left[\eta_I(x_I(t), \bar{x}_{I,P}(t), u_I(t), t) + \Delta g_{I,P}(t)\right] - L_I \epsilon_{\xi_I}(t), \\
&\Delta g_{I,P}(t) \triangleq f_I\big(C_I x_I(t), \bar{C}_I \bar{x}_{I,P}(t), u_I(t)\big) - f_I\big(C_I x_I(t) + \xi_I(t), \bar{C}_I \bar{x}_{I,P}(t) + \bar{\xi}_I(t), u_I(t)\big).
\end{aligned}
$$

Now, the residual $r_I(t)$ is "contaminated" with the fault effects that occurred in subsystem Σ_J, and we need to investigate whether this residual is bounded or not by the detection threshold $\bar{r}_I(t)$ given by (6.9) that is used by the detection agent $\mathcal{F}_I$. Following the same mathematical calculations as in the derivation of (6.23), the residual $r_I^{(j)}$ satisfies

$$
|r_I^{(j)}(t)| \leq \sum_{k=0}^{t-1} \alpha_{I,j} \delta_{I,j}^{t-1-k} \|\chi_{I,P}(k)\| + \bar{\epsilon}_{\xi_I}(t), \tag{6.32}
$$

where

$$
\begin{aligned}
\|\chi_{I,P}(t)\| &\leq \bar{H}(z)\big[\|\eta_I(x_I(t), \bar{x}_{I,P}(t), u_I(t), t)\| + \|\Delta g_{I,P}(t)\|\big] + \|L_I\| \bar{\epsilon}_{\xi_I}(t) \\
&\leq \bar{H}(z)\big[\bar{\eta}_I(y_I(t), \bar{y}_{I,P}(t), u_I(t)) + \bar{\Delta} f_I\big] + \|L_I\| \bar{\epsilon}_{\xi_I}(t).
\end{aligned} \tag{6.33}
$$

The last inequality is derived by using $\|\eta_I(x_I, \bar{x}_{I,P}, u_I, t)\| \leq \bar{\eta}_I(y_I(t), \bar{y}_{I,P}(t), u_I(t))$ (see Assumption 6.3) and $\|\Delta g_{I,P}(t)\| \leq \bar{\Delta} f_I$ (see Assumption 6.1 and (6.11)). Note that, Assumption 6.3 is stated for sensor fault-free operation because the bound on the modeling uncertainty $\bar{\eta}_I$ makes use of the measurements. In the event of a process fault which changes the state variables, Assumption 6.3 is still valid since the measurements are essentially these altered state variables (or linear combination, contaminated with the process fault effects) but with some uncertainty due to the measurement noise. In addition, note that the right side of

Christodoulos Kellis

(6.33) is actually the term $\bar{\chi}_I(t)$ given by (6.10) which is used by the the local fault detection agent $\mathcal{F}_I$ and hence $\|\chi_{I,P}(t)\| \leq \bar{\chi}_I(t)$. Therefore, from (6.32), it can be seen that the residual still satisfies $|r_I^{(j)}(t)| \leq \bar{r}_I^{(j)}(t)$ for all $j = 1, \ldots, p_I$ and hence the fault is not detected by $\mathcal{F}_I$. In other words, a process fault that occurs in a subsystem can only be detected by its respective fault detection agent.

(b) In this part, we consider fault propagation in the case of a sensor fault that occurs in Σ_J that affects Σ_I. As in the case of a process fault, the local fault detection agent $\mathcal{F}_J$ may detect the fault, but here we investigate the possibility of fault detection by the detection agent $\mathcal{F}_I$. In this case the sensor faults in Σ_J are propagated to $\mathcal{F}_I$ through the measurements $\bar{y}_I$ (communicated by $\mathcal{F}_J$, see (6.4)). For easier visual indication of the measurements $\bar{y}_I$ that contain the sensor faults, we denote them as $\bar{y}_{I,S}$. Therefore, in the case of a sensor fault occurring in Σ_J which affects Σ_I, the dynamics of Σ_I remain unaffected by the sensor fault and are given by (6.1), (6.2) whereas the estimation model of the local fault detection agent $\mathcal{F}_I$ is affected by the sensor fault and it is given by (6.4), (6.5) where now $\bar{y}_I$ is indicated with $\bar{y}_{I,S}$ (only (6.4) is affected). Note that $\bar{y}_{I,S} = \bar{y}_I^0(t) + \bar{\xi}_I(t) + \bar{\beta}_I^y(t - T_0^y)\bar{\theta}_I(t)$, where the term $\bar{\beta}_I^y(t - T_0^y)\bar{\theta}_I(t)$ indicates the sensor faults that occur in neighboring subsystems and affect Σ_I. Hence, the filtered estimation error still satisfies (6.17)-(6.19) (the fault effects enter implicitly through the measurements of the interconnection variables), which are rewritten below with explicit indication of the fault effect through $\bar{y}_{I,S}$:

$$
\begin{aligned}
&\tilde{x}_{I,f}(t+1) = A_{I,0}\tilde{x}_{I,f}(t) + \chi_{I,S}(t)\\
&\chi_{I,S}(t) \triangleq H(z)\left[\eta_I(x_I(t), \bar{x}_I(t), u_I(t), t) + \Delta g_{I,S}(t)\right] - L_I\epsilon_{\xi_I}(t),\\
&\Delta g_{I,S}(t) \triangleq f_I\big(y_I^0(t), \bar{y}_I^0(t), u_I(t)\big) - f_I\big(y_I^0(t) + \xi_I(t), \bar{y}_{I,S}(t), u_I(t)\big).
\end{aligned}
$$

Similarly to the derivation of (6.23), the residual $r_I^{(j)}$ satisfies

$$
|r_I^{(j)}(t)| \leq \sum_{k=0}^{t-1} \alpha_{I,j}\delta_{I,j}^{t-1-k}\|\chi_{I,S}(k)\| + \bar{\epsilon}_{\xi_I}(t),
$$

where

$$
\|\chi_{I,S}(t)\| \leq \bar{H}(z)\big[\|\eta_I(x_I(t), \bar{x}_I(t), u_I(t), t)\| + \|\Delta g_{I,S}(t)\|\big] + \|L_I\|\bar{\epsilon}_{\xi_I}(t).
$$

In this case, it cannot be guaranteed that $\|\eta_I(x_I, \bar{x}_I, u_I, t)\| \leq \bar{\eta}_I(y_I(t), \bar{y}_{I,S}(t), u_I(t))$ (since Assumption 6.3 might not hold due to the sensor fault) or $\|\Delta g_{I,S}(t)\| \leq \bar{\Delta} f_I$ hold and as a result $\|\chi_{I,S}(t)\|$ may exceed $\bar{\chi}_I(t)$ given by (6.10) which is used by the local fault detection agent $\mathcal{F}_I$ (note that in (6.10) the first term is actually $\bar{\eta}_I(y_I(t), \bar{y}_{I,S}(t), u_I(t))$ due to the use

Christodoulos Keliris

of the faulty measurements). Therefore, the residual may exceed its corresponding detection threshold, i.e. $|r_I^{(j)}(t)| > \bar{r}_I^{(j)}(t)$ for some $j = 1, \ldots, p_I$ which means that the sensor fault that occurred in Σ_J can be detected by the local fault detection agent $\mathcal{F}_I$. □

Remark 6.4. *A qualitative explanation can be given for Lemma 6.2 as follows. In the case of a process fault that occurs in Σ_J, the fault affects its states which in turn affect other subsystems through the interconnection variables. So, the states of Σ_J are "contaminated" by the process fault and the measurements of (some of) these states also contain the process fault effects. Therefore, a subsystem Σ_I that is affected by Σ_J, is affected by the process fault that occurred in Σ_J through the interconnection variables $\bar{C}_I\bar{x}_I$ and the detection agent $\mathcal{F}_I$ makes use of the measurements $\bar{y}_I$ which are also "contaminated" by the same fault. Hence, the effect of the process fault that occurred in Σ_J, is "canceled out" in the detection agent $\mathcal{F}_I$ and it is unable detect the fault. Hence, a process fault occurring in subsystem Σ_J is detectable only by its respective detection agent $\mathcal{F}_J$ and not by any other detection agent $\mathcal{F}_I$. On the other hand, assume that a fault is detected by the detection agent $\mathcal{F}_I$ and we know that it is a sensor fault. Then the faulty sensor might be due to the measurements y_I of Σ_I or due to the interconnection measurements $\bar{y}_I$ of the other subsystems. This is because a sensor fault occurring in one subsystem affects the estimation model of its respective detection agent $\mathcal{F}_I$ through y_I and also the estimation models of other agents $\mathcal{F}_J$ through the communicated interconnection measurements $\bar{y}_J$, whereas the actual subsystems are influenced by the fault-free states (and not the faulty measurements).*

Theorem 6.2 established a detectability condition for fault detection by the agent $\mathcal{F}_I$ when a sensor fault occurs in Σ_I. The following Theorem, gives a detectability condition for fault detection by the agent $\mathcal{F}_I$ when a sensor fault occurs in Σ_J which affects Σ_I.

Theorem 6.3 (Propagation Sensor Fault Detectability)**.** *Consider the nonlinear interconnected system (6.1), (6.2) with the distributed fault detection scheme described in (6.4), (6.5), (6.6), (6.10) and (6.9). A sensor fault that occurs at $t = T_0^y$ in Σ_J which affects Σ_I, is detectable by the local fault detection agent $\mathcal{F}_I$ if the following mismatch function $\Delta f_I''$ due to the sensor fault*

$$\Delta f_I''(t) \triangleq f_I(y_I^0(t), \bar{y}_I^0(t), u_I(t)) - f_I(y_I^0(t) + \xi_I(t), \bar{y}_I^0(t) + \bar{\xi}_I(t) + \bar{\beta}_I^y(t - T_0^y)\bar{\theta}_I(t), u_I(t))$$

Christodoulos Keliris

satisfies the following inequality at some time t, for some $j = 1, 2, \ldots, p_I$:

$$\left|\sum_{k=0}^{t-1} C_I^{(j)} A_{I,0}^{t-1-k} H(z)[\Delta f_I''(k)]\right| > \bar{r}_I^{(j)}(t) + \bar{\epsilon}_{\xi_I}(t) + \sum_{k=0}^{t-1} \alpha_{I,j}\delta_{I,j}^{t-1-k}\big(\|H(z)[\eta_I(x_I(k), \bar{x}_I(k), u_I(k), k)]\| + \|L_I\|\bar{\epsilon}_{\xi_I}(k)\big). \tag{6.34}$$

Proof. In the presence of a sensor fault that occurs at $t = T_0^y$ in Σ_J which affects Σ_I, the state estimation error (6.20) becomes

$$\tilde{x}_{I,f}(t) = A_{I,0}^t \tilde{x}_{I,f}(0) + \sum_{k=0}^{t-1} A_{I,0}^{t-1-k} \chi_I''(k), \tag{6.35}$$

where

$$\chi_I''(t) \triangleq H(z)\big[\eta_I(x_I(t), \bar{x}_I(t), u_I(t), t) + \Delta f_I''(t)\big] - L_I \epsilon_{\xi_I}(t).$$

After the occurrence of a sensor fault in Σ_J, the residual (6.6) becomes $r_I(t) = C_I \tilde{x}_{I,f}(t) + \epsilon_{\xi_I}(t)$ and by using (6.35) with $\tilde{x}_{I,f}(0) = 0$ due to the filters' initial condition, the residual is written as:

$$\begin{aligned} r_I(t) &= \sum_{k=0}^{t-1} C_I A_{I,0}^{t-1-k} \chi_I''(k) + \epsilon_{\xi_I}(t) \\ &= \sum_{k=0}^{t-1} C_I A_{I,0}^{t-1-k} H(z)[\Delta f_I''(k)] + \epsilon_{\xi_I}(t) \\ &\quad + \sum_{k=0}^{t-1} C_I A_{I,0}^{t-1-k}\big(H(z)[\eta_I(x_I(k), \bar{x}_I(k), u_I(k), k)] - L_I \epsilon_{\xi_I}(k)\big) \end{aligned}$$

By using the triangle inequality, the j-th element of $r_I(t)$ for $t > T_0^y$ satisfies:

$$\begin{aligned} |r_I^{(j)}(t)| &\geq \left|\sum_{k=0}^{t-1} C_I^{(j)} A_{I,0}^{t-1-k} H(z)[\Delta f_I''(k)]\right| - |\epsilon_{\xi_I}^{(j)}(t)| \\ &\quad - \left|\sum_{k=0}^{t-1} C_I^{(j)} A_{I,0}^{t-1-k}\big(H(z)[\eta_I(x_I(k), \bar{x}_I(k), u_I(k), k)] - L_I \epsilon_{\xi_I}(k)\big)\right| \\ &\geq \left|\sum_{k=0}^{t-1} C_I^{(j)} A_{I,0}^{t-1-k} H(z)[\Delta f_I''(k)]\right| - \bar{\epsilon}_{\xi_I}(t) \\ &\quad - \sum_{k=0}^{t-1} \alpha_{I,j}\delta_{I,j}^{t-1-k}\big(\|H(z)[\eta_I(x_I(k), \bar{x}_I(k), u_I(k), k)]\| + \|L_I\|\bar{\epsilon}_{\xi_I}(k)\big). \end{aligned}$$

For fault detection, the inequality $|r_I^{(j)}(t)| > \bar{r}_I^{(j)}(t)$ must hold for some $j = 1, 2, \ldots, p_I$, so the final fault detectability condition given in (6.34) is obtained. □

Christodoulos Keliris

As it was shown in this Section, when a process fault occurs in subsystem Σ_I, then it can only be detected by its respective fault detection agent $\mathcal{F}_I$ whereas when a sensor fault occurs in subsystem Σ_I, then it can be detected either by its respective detection agent $\mathcal{F}_I$ or any other interconnected agent $\mathcal{F}_J$ (monitoring Σ_J which is affected by Σ_I). This discriminating factor is exploited in the following Section for devising a high-level fault isolation scheme.

6.5 High-level fault isolation

In this section, we exploit the findings of the previous analysis for the derivation of a high-level fault isolation scheme (see Figure 6.1). The purpose of this high-level isolation scheme is to infer some conclusions regarding the type and/or location of the fault that has occurred in the whole interconnected system according to the decisions of the detection agents, although, it does not necessarily mean that exact fault identification/isolation can be achieved. As a result, the high-level isolation scheme can provide valuable information that can be used by a more advanced fault isolation scheme in order to greatly improve its performance by excluding potential fault scenarios.

In the analysis so far, we have considered the cases of process and sensor fault separately in order to gain some intuition of how these faults affect the local and neighboring detection agents. The main conclusion of this analysis was given in Lemma 6.2 which also constitutes the basis of the subsequent high-level isolation scheme. Although the proposed fault detection scheme may handle multiple faults, which can be both process and sensor faults, for the sake of the proposed high-level isolation scheme it is assumed that only one fault can occur among all subsystems, which may either be a process or a single sensor fault. Ideally, we would like to identify the type of fault that has occurred, that is, whether it is a process or sensor fault and furthermore in the case of a sensor fault to identify the faulty sensor.

In order to identify the sensors let us consider the following:

Definition 6.1. *Let $\mathcal{S}\{y_I\}$ be the set of p_I sensors that measure $y_I \in \mathbb{R}^{p_I}$, $\mathcal{S}\{y_I\} \cap \mathcal{S}\{y_J\}$ be the set of common sensors among y_I and y_J and $\mathcal{S}\{y_I\} \cup \mathcal{S}\{y_J\}$ be the union of the set of sensors that measure y_I and y_J.*

In addition let $\mathcal{M}$ be the set of indices of the local fault detection agents that have detected a fault, i.e. $\mathcal{M} \triangleq \{I \in \{1, \ldots, N\} : \mathcal{F}_I \text{ detects fault}\}$, let $\mathcal{M}_i$ indicate the i-th index of the set $\mathcal{M}$ and, let $m \in \{0, 1, \ldots, N\}$ be the cardinality of the set $\mathcal{M}$, i.e. $m \triangleq card(\mathcal{M})$.

Then the following high-level isolation facts can be deduced:

Christodoulos Kallis

- If $m = 0$ then all subsystems are considered as potentially non-faulty (the possibility of a fault not yet detected cannot be excluded).

- If $m = 1$ then the fault may be one of the following:

  a) a process fault that has occurred in $\Sigma_{\mathcal{M}_1}$, OR

  b) a single sensor fault in $\mathcal{S}\{y_{\mathcal{M}_1}\}$, OR

  c) a single sensor fault in $\mathcal{S}\{\bar{y}_{\mathcal{M}_1}\}$.

- If $m \geq 2$ then the possibility of a process fault can be excluded and hence it is guaranteed that a sensor fault has occurred. The faulty sensor can be isolated within the set given by

$$\bigcap_{i=1}^{m} \Big(\mathcal{S}\{y_{\mathcal{M}_i}\} \cup \mathcal{S}\{\bar{y}_{\mathcal{M}_i}\} \Big). \tag{6.36}$$

If we consider that multiple faults can occur in the subsystems (i.e. multiple process and/or multiple sensor faults) then the isolation logic is modified as follows:

- If $m = 0$ then all subsystems are considered as potentially non-faulty.

- If $m \geq 1$ then the fault(s) may be:

  a) a single/multiple process fault(s) that occurred in $\Sigma_{\mathcal{M}_i}, i \in \{1, \ldots, m\}$, AND/OR

  b) a single/multiple sensor fault(s) within the set given by

$$\bigcup_{i=1}^{m} \Big(\mathcal{S}\{y_{\mathcal{M}_i}\} \cup \mathcal{S}\{\bar{y}_{\mathcal{M}_i}\} \Big). \tag{6.37}$$

The proposed high-level fault isolation scheme in the case of joint process and sensor faults assuming that only one fault can occur, results in more constrained fault possibilities in comparison to the multiple fault case, in the sense that if two or more fault detection agents detect a fault then the occurrence of a process fault is excluded, and hence the fault is guaranteed to be a sensor fault contained in the set of sensors described by (6.36). If only one fault detection agent detects a fault then the fault type cannot be determined apart from the fact that the fault can either be a process or a sensor fault in the agents' respective subsystem, or a sensor fault in neighboring subsystems that affect the subsystem the fault has been detected in. In the case we consider multiple faults, the isolation results include more possibilities about fault occurrences but still provide useful information about the fault type and the set of that the faulty sensors are contained in.

Christodoulos Kellris

These results prove to be more valuable in the case we consider that only one type of fault can occur (process or sensor), as it is the majority of the research conducted in the literature. By considering that only process faults can occur, that is by assuming that all sensors are healthy, then the detection of a fault by any fault detection agent guarantees that the fault has occurred in the respective subsystem that the particular agent is monitoring. Hence, partial fault isolation is achieved in the sense that the faulty subsystem is identified. Even if two or more agents detect a fault, then this means that in each of the respective subsystems a process fault has occurred. In the case we consider only sensor faults can occur (no process faults), then the faulty sensor(s) can be isolated within the set given by (6.36) in the case of a single sensor fault or by (6.37) in the case of multiple sensor faults. This additional information can be used by a more sophisticated fault isolation scheme in order to enhance its performance, by excluding potential fault scenarios. Moreover, the proposed distributed fault detection approach has significant benefits in comparison to a centralized approach since it encompasses important fault isolation characteristics.

In the proposed scheme, each detection agent provides a binary decision regarding the detection of a fault in the subsystem it monitors and, according to the decisions of all the agents, the high-level isolation scheme provides some information regarding the type and location of the fault that occurred. Moreover, some hypotheses can be stated regarding the status of presence of the fault, i.e. if the fault is permanent, intermittent or temporary. For instance, the successive threshold crossings, in the event of fault detection, can be interpreted as "the fault(s) is still present" but, the scheme cannot determine precisely the status of the fault. Of course, some hypotheses can be stated; according to the rate of successive threshold crossings and the frequency this behavior is observed. For instance, permanent faults, given that they are sufficiently large, will most probably cause the residuals to exceed their thresholds almost all the time after the initial detection or at least demonstrate rapid threshold crossings. On the other hand, temporary faults can be identified if after some time of the initial fault detection, the residuals fall and stay below their corresponding thresholds indicating potentially healthy operation. Finally, intermittent faults behavior, will exhibit a mixture of the permanent and temporary faults behavior, with these two phases repeating successively. Specifically, during a period of time the residual will exceed its threshold like in the case of a permanent fault, and afterwards it will be followed by a period of time where the residuals stay below their thresholds. In general, the issue of distinguishing between permanent, temporary and intermittent faults, requires further investigation and is out of the scope of the present work.

Christodoulos Kallis

6.6 Simulation Results

In this section, we consider a numerical example based on a system of two inverted pendulums connected by a spring. The discrete time models of the two subsystems $I = 1, 2$ are obtained from a modified version of the continuous time version in [78] by using a forward Euler discretization with a time step $T_s = 0.0001$s and are given by

$$\begin{aligned} x_I^{(1)}(t+1) =& x_I^{(1)}(t) + T_s x_I^{(2)}(t) \\ x_I^{(2)}(t+1) =& x_I^{(2)}(t) + T_s \left(f_I^{(2)}(t) + w_I^{(2)}(t) + \eta_I^{(2)}(t) \right) \\ y_I(t) =& x_I^{(1)}(t) + \xi_I(t) \end{aligned}$$

where for the first subsystem the nominal and interconnection functions are given by:

$$\begin{aligned} f_1^{(2)}(t) =& \left(\frac{m_1 g r}{J_1} - \frac{k r^2}{4 J_1} \right) \sin(x_1^{(1)}(t)) + \frac{kr}{2J_1}(l-b) + \frac{u_1}{J_1} \\ w_1^{(2)}(t) =& \frac{kr^2}{4J_1} \sin(x_2^{(1)}(t)) \end{aligned}$$

and for the second subsystem the respective functions are

$$\begin{aligned} f_2^{(2)}(t) =& \left(\frac{m_2 g r}{J_2} - \frac{k r^2}{4 J_2} \right) \sin(x_2^{(1)}(t)) + \frac{kr}{2J_2}(l-b) + \frac{u_2}{J_2} \\ w_2^{(2)}(t) =& \frac{kr^2}{4J_2} \sin(x_1^{(1)}(t)). \end{aligned}$$

The modification of this model with respect to [78], is with regards to the availability of the state variables for measurement, the presence of modeling uncertainty and measurement noise. Specifically, in [78] full state measurement is considered, whereas in this example it is considered that only $x_I^{(1)}(t)$, $I = 1, 2$ can be measured (with some uncertainty).

The parameters that are used in the simulation are: m_1=2kg, m_2=2.5kg, J_1=0.5kg, J_2=0.625kg, k=30N/m, l=0.5m, b=0.4m and g=9.81m/s^2. The modeling uncertainties of the subsystems are assumed to be $\eta_I^{(2)}(x_I^{(1)}, t) = 0.05\sin(10t) + 0.05 sin(x_I^{(1)}(t))$, $I = 1, 2$, in which the term $0.05\sin(10t)$ corresponds to the uncertainty associated with time variations or general inaccuracies, whereas the term $0.05 sin(x_I^{(1)}(t))$ corresponds to the uncertainty of the nominal function due to the error on some model parameters (i.e. the mass). Note that, the modeling uncertainty is a function of both the time t and the state $x_I^{(1)}$. The bound on the modeling uncertainty that is used is given by $\bar{\eta}_I^{(2)}(y_I^{(1)}, t) = 0.1 + 0.05|sin(y_I^{(1)}(t))|$, $I = 1, 2$. The inputs u_I are derived based on a simple decentralized proportional feedback controller that stabilizes each subsystem and are given by $u_I = 20 e_I$, $I = 1, 2$ where $e_I = -y_I^{(1)}$ is the tracking error. In this example, we consider two cases, one for process fault and one for the

Christodoulos Keliris

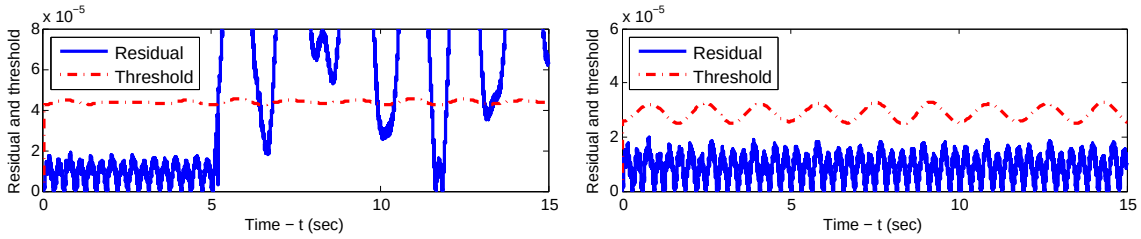
sensor fault. In the case of a process fault, we consider an abrupt multiplicative actuator fault in subsystem 1 where the input changes to $u_1 = (1 + \beta_1)\bar{u}_1$, where $\bar{u}_1$ is the nominal control input in the non-fault case and $\beta_1 \in [-1, 0]$ is the parameter characterizing the magnitude of the fault. The actuator fault in this case can be considered as a process fault affecting the dynamics of the system. The fault occurs at $T_0^x = 5$ sec with a magnitude $\beta_1 = -0.1$. In the case of a sensor fault, we consider that the sensor in the first subsystem measuring y_1 measures the signals' amplitude with 20% deviation and the sensor fault occurs at $T_0^y = 5$ sec.

The measurement noise ξ_I is implemented as a uniform random number in the range $[-0.01, 0.01]$. The proposed fault detection scheme is implemented using a FIR filter for $H(z)$. Specifically, the filter $H(z)$ is designed as a 10-th order FIR lowpass filter with normalized cutoff frequency 0.2 and utilizing a Hamming window (using the `fir1` command in Matlab). The transfer function of $H(z)$ is given by $H(z) = \sum_{k=0}^{10} d_k z^{-k}$ and explained in Section 6.2.3 the filter $\bar{H}(z)$ is given by $\bar{H}(z) = \sum_{k=0}^{10} |d_k| z^{-k}$. Using the aforementioned filter $H(z)$ the bounds on the filtered noise are found through the simulation as $\bar{\epsilon}_{\xi_I} = 4\text{e}{-6}$ and $\bar{\epsilon}_{\Delta f_I} = 1.2\text{e}{-4}$ for $I = 1, 2$.

According to the proposed fault detection scheme, two fault detection agents are designed, one for each subsystem. In each detection agent, the estimation model is given by (6.4), (6.5), the residual is generated according to (6.6) and finally, the detection threshold is generated according to (6.9) with (6.27) to fully exploit the filtering benefits. For the threshold implementation the constants $\alpha_{I,1} = 1, \delta_{I,1} = 0.7$, $I = 1, 2$, are also used so that (6.12) is satisfied. The simulation results for the fault detection agents that monitor the first subsystem (measured variable y_1) and the second subsystem (measured variable y_2) in the case of the process fault are shown in Figure 6.3 and in the case of the sensor fault in Figure 6.5.

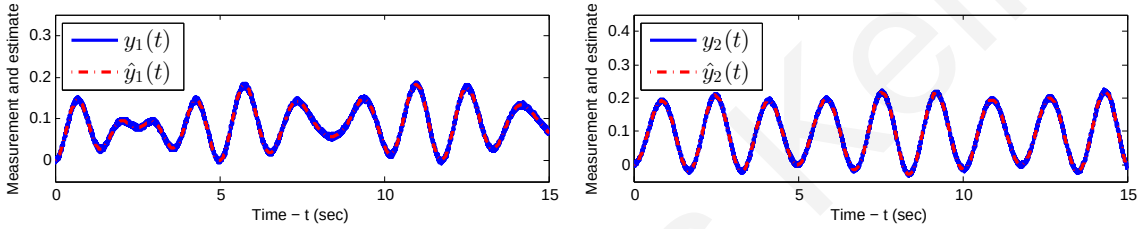
At first let's consider the case of the process fault. In this case the measurements and their corresponding estimates of the two subsystems are shown in Figure 6.4 where it can be seen that the estimation model tracks the measurement although, it cannot be seen any significant discrepancy due to the process fault. The discrepancy of course is present, and it is the cause for the residual signal given in Figure 6.3a to exceed its threshold. Specifically, the simulation results that correspond to the detection agent that monitors the measurements y_1 (1st subsystem) are shown in Figure 6.3a where it is clearly seen that the residual significantly exceeds its detection threshold after the fault occurs, hence the fault is detected at around $t = 5.09$ sec. It must be noted that there are no false alarms prior to the fault occurrence in either case. The corresponding results of the detection agent that monitors y_2 (2nd subsystem) are

Christodoulos Keliris



(a) Fault detection agent monitoring measurement y_1. (b) Fault detection agent monitoring measurement y_2.

Figure 6.3: Residual signal and fault detection threshold for measurements y_1, y_2 in the case of process fault occurring in Σ_1.



(a) Measurement y_1 and estimate $\hat{y}_1$.

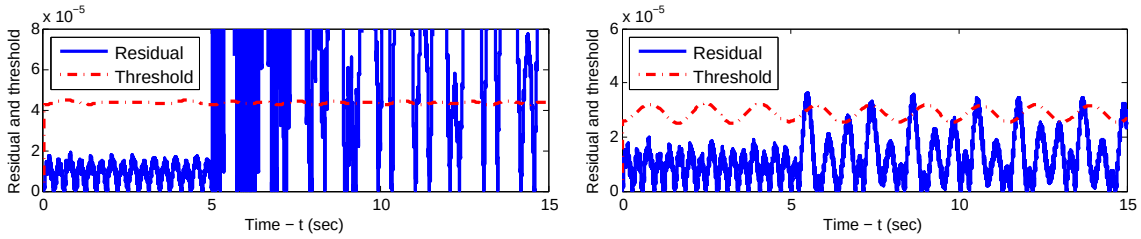
(b) Measurement y_2 and estimate $\hat{y}_2$.

Figure 6.4: Measurements y_1, y_2 and their corresponding estimates $\hat{y}_1$, $\hat{y}_2$ in the case of process fault occurring in Σ_1.

shown in Figure 6.3b where it is seen that the residual signal is always below its threshold signal and therefore no fault is detected. It must be noted that, the residual in the case of the second agent monitoring y_2 in Figure 6.3b, exhibits the same behavior before and after the occurrence of the fault, indicating that the process fault effects from the first subsystem do not seem to impact the residual of the second agent.

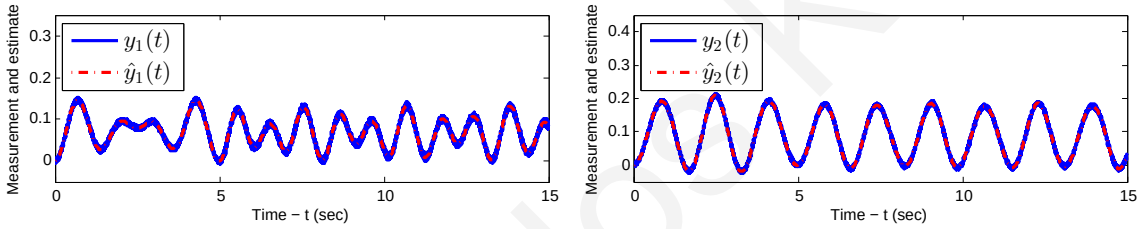
Now let us consider the case of the sensor fault. In this case, the measurements and their corresponding estimates of the two subsystems are shown in Figure 6.6. In the case of the sensor fault, the results in the case of the fault detection agents that monitor the measurements y_1 and y_2 are shown in Figure 6.5a and 6.5b respectively. In this case, the sensor fault is detected very fast by the detection module that monitors y_1 at around $t = 5.01$ sec. Most importantly though, the sensor fault is also detected by the detection module that monitors y_2 at around $t = 5.39$ sec. As in the previous case, no false alarms occur, since the residuals are always bounded by their thresholds prior to the sensor fault occurrence. In this case, it must be noted that, the residual behavior in the case of the second agent monitoring y_2 in Figure 6.5b, changes after the occurrence of the fault, indicating that the sensor fault effects from the first subsystem affect significantly the residual of the second agent.

Christoforos Keliris



(a) Fault detection agent monitoring measurement y_1. (b) Fault detection agent monitoring measurement y_2.

Figure 6.5: Residual signal and fault detection threshold for measurements y_1, y_2 in the case of sensor fault occurring in Σ_1.



(a) Measurement y_1 and estimate $\hat{y}_1$. (b) Measurement y_2 and estimate $\hat{y}_2$.

Figure 6.6: Measurements y_1, y_2 and their corresponding estimates $\hat{y}_1$, $\hat{y}_2$ in the case of sensor fault occurring in Σ_1.

The simulation results are in line with the findings of the analysis conducted in Section 6.3 and more specifically with Lemma 6.2 according to which the process fault occurring in a subsystem can only be detected by its respective fault detection agent, whereas a sensor fault occurring in a subsystem can be detected by its respective fault detection and also by neighboring interconnected detection agents. According to the high-level fault isolation scheme described in Section 6.5, in the case of the process fault, no conclusive decision can be reached since only one detection agent detects a fault and therefore the fault can be either a process fault in Σ_1 (which is actually the case), or a fault in the sensors measuring y_1 or y_2. In the case of the sensor fault, because both fault detection agents detect a fault, the case of a process fault is excluded and hence it is guaranteed that a sensor fault has occurred and the faulty sensor is either y_1 (which is actually the case) or y_2. Therefore, further actions can be taken to identify precisely the type of the fault that has occurred.

Christodoulos Keliris

6.7 Conclusion

In this chapter, a distributed fault diagnosis approach for the detection of process and sensor faults in a class of interconnected input-output discrete-time, nonlinear systems with modeling uncertainties and measurement noise is presented. By utilizing a filtering approach which is incorporated in the fault detection framework to mitigate the measurement noise effects, robust adaptive thresholds are designed that guarantee no false alarms. Furthermore, the propagation of a fault that occurs in one subsystem and affects neighboring subsystems is investigated, leading to some key properties of fault propagation among subsystems. More specifically, the fault detection scheme is designed in such a way that a process fault occurring in a subsystem can only be detected by its corresponding detection agent, whereas a sensor fault occurring in a subsystem can also be detected by the detection agents of the neighboring subsystems it affects. This discriminating element is exploited to extrapolate further information regarding the type of fault that has occurred and, constitutes the basis of the derived high-level isolation scheme. Furthermore, detectability conditions have been derived that characterize quantitatively the class of process and sensor faults that can be detected by the proposed scheme.

Christodoulos Kellis

Chapter 7

An Integrated Learning and Filtering Approach for Fault Diagnosis

This chapter develops an integrated filtering and adaptive approximation based approach for fault diagnosis of process and sensor faults in a class of continuous-time nonlinear systems with modeling uncertainties and measurement noise. The proposed approach utilizes the designed filtering framework described in Chapter 3 for dampening the measurement noise effect and incorporates learning in order to learn the modeling uncertainty, enhance the fault detectability and allow the discrimination between process and sensor faults. The primary objective and main contribution of this chapter is the design of a unified fault diagnosis approach by: a) integrating learning and filtering techniques for obtaining tight detection thresholds that guarantee no false alarms and thus, enhancing fault detectability and, b) identifying the fault type (process or sensor) when a fault is detected and generating an estimation of the fault. The first objective is achieved in two ways; i) at first, adaptive approximation methods are used for learning the modeling uncertainty (so that the learned modeling uncertainty function is used in the design of the residual signals) and then ii) by using filtering to attenuate the effect of measurement noise on the diagnosis thresholds. Both these tasks are integrated in a unified framework and intertwined through the filtering approach devised in Chapter 3, which is now decomposed in a two stage filtering process in order to derive the required signals for the adaptive approximation and for the residual derivation. For achieving the second objective (fault type identification), two estimation models are constructed, one for process and one for sensor faults, that utilize adaptive approximation methods to learn the potential faults and the fault type is determined on an exclusion-based logic. Finally, fault detectability and identification conditions are derived for both process and sensor faults, that

Christodoulos Kellris

provide an implicit characterization of the faults that can be detected and identified by the proposed scheme.

The chapter is organized as follows: in Section 7.1 the problem formulation is given, in Section 7.2 the detailed design of the fault detection scheme by combining adaptive approximation with filtering is presented in detail and, in Section 7.3 the detectability conditions for process and sensor faults that characterize the class of detectable faults are derived. The details of the fault identification procedure are given in Section 7.4 and, in Section 7.5 the fault-type identification conditions are derived. In Section 7.6 a simulation example demonstrating the effectiveness of the scheme is presented and, finally, in Section 7.7 some concluding remarks are provided.

7.1 Problem Formulation

Consider a nonlinear dynamic system:

$$\Sigma : \begin{cases} \dot{x}(t) = Ax(t) + f(x(t), u(t)) + \eta(x(t), u(t)) + \beta^x(t - T_0^x)\phi(x(t), u(t)) & (7.1) \\ y(t) = x(t) + \xi(t) + \beta^y(t - T_0^y)\sigma(t), & (7.2) \end{cases}$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$ are the state, input and measured output vectors respectively, the matrix $A \in \mathbb{R}^{n\times n}$ and the function $f : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^n$ are the known (nominal) function dynamics and $\eta : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^n$ is the modeling uncertainty associated with the nominal function. The vector $\xi \in \mathcal{D}_\xi \subset \mathbb{R}^n$ ($\mathcal{D}_\xi$ is a compact set) represents the unknown measurement noise. The term $\beta^x(t - T_0^x)\phi(x, u)$ characterizes the time-varying process fault function dynamics affecting the system. More specifically, the term $\phi : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^n$ represents the unknown process fault function and the term $\beta^x(t - T_0^x) : \mathbb{R} \mapsto \mathbb{R}^+$ models the time evolution of the process fault, where T_0^x is the unknown time of the process fault occurrence. The time profile $\beta^x(t - T_0^x)$ can be used to model both abrupt and incipient faults. Specifically, prior to the fault occurrence the time profile is considered to be zero, i.e. $\beta^x(t - T_0^x) = 0$, for all $t < T_0^x$, and after the fault occurrence it is monotonically increasing to one as $t \to \infty$ (in the case of incipient fault, since in the case of abrupt faults the time profile takes the form of a step function). The term $\beta^y(t - T_0^y)\sigma(t)$ characterizes the time-varying magnitude of the sensor faults. More specifically, each component $\sigma^{(k)}(t)$, $k = 1, 2, \ldots, n$, of the sensor fault vector $\sigma : \mathbb{R}^+ \mapsto \mathbb{R}^n$, represents the (possibly time varying) bias due to a sensor fault that occurs in the k-th sensor at time T_0^y. In this work, we consider only the case

Christodoulos Kellis

of abrupt sensor faults, i.e.

$$\beta^y(t - T_0^y) = \begin{cases} 0 \text{ if } t < T_0^y \\ 1 \text{ if } t \geq T_0^y. \end{cases} \tag{7.3}$$

Please note that, without loss of generality, in this chapter all the sensor faults are considered to occur at time T_0^y for simplicity. In addition, we consider the occurrence of a single type of fault: either a process fault occurring at time T_0^x or single/multiple sensor fault(s) occurring at time T_0^y.

The objective is twofold: i) to exploit the filtering framework, developed in Chapter 3, in order to exploit the noise suppression properties of filtering and to integrate learning methods for approximating the modeling uncertainty so that tight detection thresholds are obtained, thus enhancing fault detectability and, ii) when a fault is detected, to determine whether it is a process or sensor fault and provide an estimation of the fault function (i.e. $\hat{\phi}, \hat{\sigma}$). The formulation is independent of the controller used, which may either be a generic one achieving some desired control objectives or one that incorporates an advanced fault accommodation scheme. The learning of the overall uncertainty function η is based on adaptive approximation methods and allows the use of the learned function dynamics (indicated by $\hat{\eta}$) in a suitable residual signal to "cancel out" the true function η and hence, aid in the derivation of a tighter detection threshold.

After the detection of a fault, two Fault Identification Estimators (FIEs) are initiated to identify the type of the fault that has occurred and provide an estimation of the fault. Each FIE corresponds to a fault type (i.e. process or sensor) and by utilizing adaptive approximation methods, the fault that has occurred is being learned whilst, at the same time, residual and isolation thresholds are derived that are used for the identification of the fault type. The fault type identification is based on an exclusion logic, in the sense that when a particular residual generated from a FIE exceeds its corresponding isolation threshold then, the particular fault type that the FIE corresponds to, is excluded. Hence, it is guaranteed that the fault type is the other one, for which all the residuals of its FIE model remain below their corresponding isolation thresholds, and the fault estimation is provided by the latter FIE model. Figure 7.1 illustrates the general procedure of the scheme developed in this chapter.

In the sequel $|\cdot|$ indicates the absolute value of a scalar function, the Euclidean 2-norm for vectors and, the matrix norm induced by the 2-norm for matrices.

***Assumption* 7.1.** *The state variable x and the local input u remain bounded in some compact region of interest $\mathcal{D} = \mathcal{D}_x \times \mathcal{D}_u \subset \mathbb{R}^n \times \mathbb{R}^m$, before and after the occurrence of a fault.*

Christodoulos Kallis

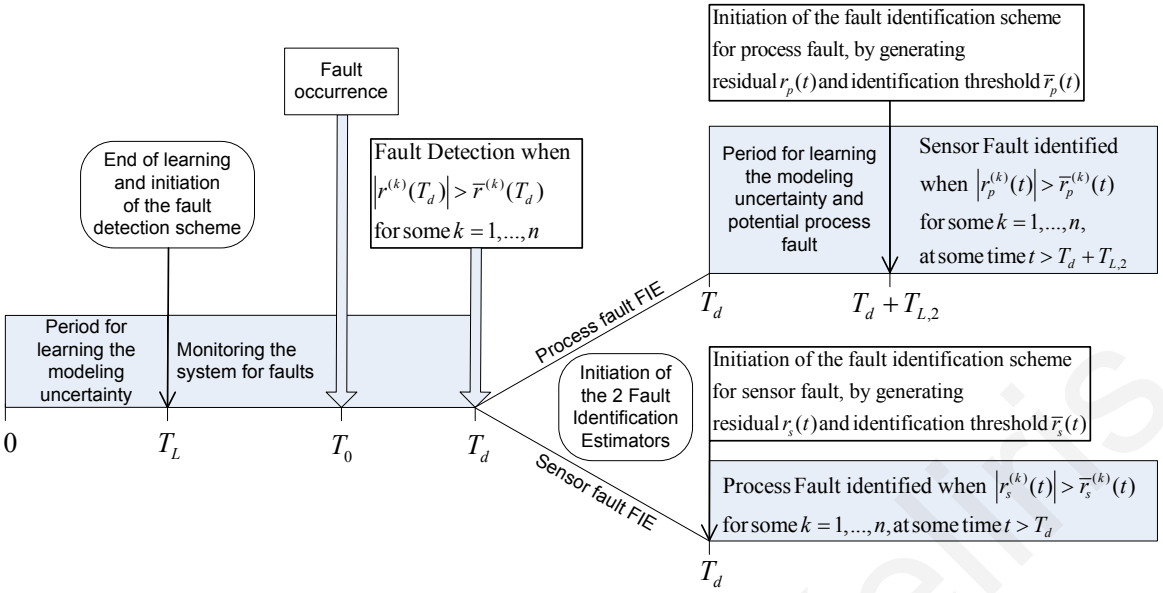


Figure 7.1: Diagram describing the general procedure of the fault detection and identification scheme. The time T_L indicates the learning time allowed for learning the modeling uncertainty, T_0 indicates the time of the fault occurrence (process or sensor), T_d indicates the fault detection time and, $T_{L,2}$ indicates the learning time allowed for learning the combined effect of the modeling uncertainty with the potential process fault.

***Assumption* 7.2.** *The k-th component of the function $f(x,u)$ satisfies the Lipschitz condition:*

$$|f^{(k)}(x_1,u) - f^{(k)}(x_2,u)| \leq \lambda_{f_k}|x_1 - x_2| \tag{7.4}$$

for $x_1, x_2 \in \mathcal{D}_x$, $u \in \mathcal{D}_u$ where λ_{f_k} is the known Lipschitz constant. Hence, for the vector function $f(x,u)$, the Lipschitz condition becomes $|f(x_1,u) - f(x_2,u)| \leq \lambda_f|x_1 - x_2|$ where λ_f is the Lipschitz constant which can be considered as $\lambda_f \triangleq \sum_{k=1}^{n} \lambda_{f_k}$.

***Assumption* 7.3.** *The rate of change of the sensor bias $\sigma^{(k)}(t)$ is unifromly bounded as follows:*

$$|\dot{\sigma}^{(k)}(t)| \leq \psi, \qquad k = 1,\ldots,n, \tag{7.5}$$

where ψ is a known positive scalar.

***Assumption* 7.4.** *The measurement noise is bounded, i.e. $|\xi^{(k)}(t)| \leq \bar{\xi}^{(k)}$, where $\bar{\xi}^{(k)}$ is a known constant. Hence, $|\xi(t)| \leq \bar{\xi}_d \triangleq |\bar{\xi}|$.*

Assumption 7.1 is required for well-posedness since in this work we address the fault detection problem, not the control design and fault accommodation problem. Assumptions 7.2

Christodoulos Keliris

and 7.3 are required for the design of the FIE model for the sensor faults and, Assumption 7.4 is required in order to distinguish the effects of sensor faults.

In the following section, the details of the proposed scheme for the tasks of learning the modeling uncertainty and designing suitable fault detection thresholds are given.

7.2 Fault Detection

First, we derive suitable detection thresholds by computing uncertainty bounds associated with the fault-free operation of the system.

7.2.1 Filtering

Each measured variable $y^{(k)}(t)$ is filtered by an induced filter $H(s)$ which is implemented as a series of two filters $H_1(s)$ and $H_2(s)$, for reasons that will become apparent in the sequel, such that $H(s) = H_1(s)H_2(s)$. The filter $H(s)$ is of p-th order with a strictly proper transfer function $H(s) = sH_p(s)$, where

$$H_p(s) = \frac{d_{p-2}s^{p-2} + d_{p-3}s^{p-3} + \ldots + d_0}{s^p + c_{p-1}s^{p-1} + \ldots + c_1 s + c_0}. \tag{7.6}$$

At first, the selection of the filter $H(s)$ is made so that the effect of the measurement noise is dampened. Then, the filters $H_1(s)$ and $H_2(s)$ are obtained by introducing the design constant $\alpha > 0$ (which is required for the learning task) and are given by:

$$H_1(s) = \frac{s}{s+\alpha} \tag{7.7}$$

$$H_2(s) = \frac{(s+\alpha)\big(d_{p-2}s^{p-2} + d_{p-3}s^{p-3} + \ldots + d_0\big)}{s^p + c_{p-1}s^{p-1} + \ldots + c_1 s + c_0}. \tag{7.8}$$

In the following, the initial conditions of the filters $H_1(s), H_2(s)$ are considered to be zero.

It must be stressed that, although four filters have been stated above ($H(s)$, $H_p(s)$, $H_1(s)$, $H_2(s)$), only $H(s)$ is required to be selected, so that the noise attenuation is achieved, and then, the other three filters can be easily deduced from the particular selection of $H(s)$.

The filters $H_1(s)$ and $H_2(s)$ (and hence $H(s)$ and $H_p(s)$) are asymptotically stable and therefore BIBO stable. Therefore, for bounded measurement noise $\xi(t)$, the filtered measurement noise $\epsilon_\xi(t) \triangleq H(s)\,[\xi(t)]$ is bounded as follows:

$$|\epsilon_\xi^{(k)}(t)| \leq \bar{\epsilon}_\xi^{(k)}(t) \quad k = 1, 2, \ldots, n,$$

where $\bar{\epsilon}_\xi^{(k)}$ are bounding functions that are computable since $\xi \in \mathcal{D}_\xi$.

Christodoulos Kelins

Generally, each measured variable $y^{(k)}(t)$ can be filtered by a different induced filter $H(s)$ with different design constants α. In this chapter, without loss of generality, we consider $H(s)$ and α to be the same for all the output variables in order to simplify the notation and presentation.

The two-step filtering process using filters $H_1(s)$ and $H_2(s)$ is employed, instead of directly filtering each measurement with $H(s)$, in order to integrate the filtering with the learning task and, be able to derive the required signals for both tasks (approximation error for learning and residual for fault detection). It must be noted that, adaptive approximation methods require that the filters used are Strictly Positive Real (SPR) [24, 43]. Therefore, the filter $H(s)$, which is selected for dampening the measurement noise, may not be SPR and hence, not suitable for the learning task. With the decomposition of $H(s)$ into the two filters $H_1(s)$ and $H_2(s)$, the output of the filter $H_1(s)$ (which is SPR) can be used for the learning task and, the output of the filter $H_2(s)$ can be used for the fault detection task. Therefore, the decoupling of the two tasks into two separate and independent ones is achieved.

7.2.2 Adaptive approximation

An adaptive approximator is used for learning the modeling uncertainty η and providing an estimate $\hat{\eta}$. To simplify the notation in the following analysis, (7.1) is rewritten as

$$\dot{x}(t) = g(x(t), u(t)) + \eta(x(t), u(t)) + \beta^x(t - T_0^x)\phi(x(t), u(t)), \tag{7.9}$$

where $g(x(t), u(t)) \triangleq Ax(t) + f(x(t), u(t))$.

Based on (7.9), an estimation model $\hat{x}(t)$ for $x(t)$ under fault-free operation is generated as follows:

$$\dot{\hat{x}}(t) = g(y(t), u(t)) + \hat{\eta}(y(t), u(t), \hat{\theta}(t)), \tag{7.10}$$

with initial condition $\hat{x}(0) = y(0)$, $\hat{\eta}$ denotes the output of an adaptive approximator structure and $\hat{\theta} \in \mathbb{R}^q$ is a set of adjustable parameters in vector form. More details regarding the adaptive approximator design will be given in the sequel.

The signal which is used in the adaptive law for adjusting the parameter vector $\hat{\theta}$ is given by

$$\epsilon(t) \triangleq H_1(s)\big[y(t) - \hat{x}(t)\big], \tag{7.11}$$

and by using (7.2), (7.7) and the Laplace differentiation property $s[x(t)] = \dot{x}(t) + x(0)\delta(t)$

Christodoulos Kallis

(where $\delta(t)$ is the delta function) we obtain

$$\begin{aligned}\epsilon(t) &= \frac{s}{s+\alpha}\big[x(t)+\xi(t)-\hat{x}(t)\big] \\ &= \frac{1}{s+\alpha}\big[\dot{x}(t)+x(0)\delta(t)\big]+\frac{s}{s+\alpha}\big[\xi(t)\big]-\frac{1}{s+\alpha}\big[\dot{\hat{x}}(t)+\hat{x}(0)\delta(t)\big]. \end{aligned} \tag{7.12}$$

By using (7.1) and (7.10), (7.12) becomes

$$\epsilon(t) = \frac{1}{s+\alpha}\big[\Delta g(t)+\Delta\eta(t)-\alpha\xi(t)\big]-e^{-\alpha t}\xi(0)+\xi(t) \tag{7.13}$$

where

$$\Delta g(t) \triangleq g\big(x(t),u(t)\big)-g\big(y(t),u(t)\big) \tag{7.14}$$

$$\Delta\eta(t) \triangleq \eta\big(x(t),u(t)\big)-\hat{\eta}\big(y(t),u(t),\hat{\theta}(t)\big). \tag{7.15}$$

As it can be seen from (7.13), $\epsilon(t)$ is comprised by the functional error $\Delta\eta(t)$ and some additional terms, which arise due to the measurement noise $\xi(t)$. Therefore, the signal $\epsilon(t)$ provides a measure of $\Delta\eta(t)$ and, by an appropriate selection of α, the scaled quantity $\alpha\epsilon(t)$ (which is measurable) provides a good approximation of $\Delta\eta(t)$ (which is unknown) over the lower frequency range (depending on the pole of the filter $H_1(s)$). Hence, $\epsilon(t)$ can be used in the adaptive law for adjusting the parameter vector $\hat{\theta}$ of the approximator [69].

In this work, we use a linearly parametrized approximator $\hat{\eta}(y,u,\hat{\theta})$ so that each component $k=1,2,\ldots,n$ is given by $\hat{\eta}^{(k)}(y,u,\hat{\theta}_k)=\omega_k^\top(y,u)\hat{\theta}_k$, where $\omega_k:\mathbb{R}^n\times\mathbb{R}^m\mapsto\mathbb{R}^{q_k}$ is a vector composed of smooth functions independent of $\hat{\theta}_k$, and $\hat{\theta}_k\in\mathbb{R}^{q_k}$ is a set of adjustable parameters in vector form. In compact form, we have

$$\hat{\eta}(y,u,\hat{\theta}) = \Omega(y,u)\hat{\theta}, \tag{7.16}$$

where $\Omega:\mathbb{R}^n\times\mathbb{R}^m\mapsto\mathbb{R}^{n\times q}$ is the diagonal block matrix $\Omega(y,u)=\omega_1^\top(y,u)\oplus\omega_2^\top(y,u)\oplus\ldots\oplus\omega_n^\top(y,u)$, where $\oplus$ indicates the direct sum operator and $\hat{\theta}\in\mathbb{R}^q$ ($q=\sum_{k=1}^n q_k$) is the column vector $\hat{\theta}=\big[\hat{\theta}_1^\top,\ldots,\hat{\theta}_n^\top\big]^\top$.

Based on (7.13) and by using techniques from adaptive control (Lyapunov synthesis approach, see [43]), $\hat{\theta}$ is updated according to the following law:

$$\dot{\hat{\theta}}(t) = \mathcal{P}_p\big(\Gamma^x\Omega^\top(t)\epsilon(t)\big), \tag{7.17}$$

where $\Gamma^x\in\mathbb{R}^{q\times q}$ is a symmetric and positive definite learning rate matrix, and $\mathcal{P}_p$ is a projection operator that restricts $\hat{\theta}$ in a predefined compact and convex set $\hat{\Theta}_p\in\mathcal{R}^q$ in order to avoid parameter drift, a phenomenon that may appear with standard adaptive laws in the presence

Christodoulos Keliris

of modeling uncertainties. Note that due to the linearized parametrized approximator setting, (7.17) is simplified to $\dot{\hat{\theta}}_k = \mathcal{P}_{p_k}\left(\Gamma_k^x \omega_k^\top(t)\epsilon^{(k)}(t)\right), \quad k = 1, 2, \ldots, n$, where $\Gamma_k^x \in \mathbb{R}^{q_k \times q_k}$ are symmetric and positive definite learning rate matrices. In this chapter, $\hat{\Theta}_p$ is considered to be an origin-centered hypersphere with radius M_p, i.e. $\hat{\Theta}_p \triangleq \left\{\hat{\theta} \in \mathcal{R}^q : |\hat{\theta}| \leq M_p\right\}$ and hence the adaptive law (7.17) with the projection algorithm can be expressed in the form:

$$\dot{\hat{\theta}}(t) = \Gamma^x \Omega^\top(t)\epsilon(t) - \iota^* \Gamma^x \frac{\hat{\theta}(t)\hat{\theta}^\top(t)}{\hat{\theta}^\top(t)\Gamma^x\hat{\theta}(t)}\Gamma^x \Omega^\top(t)\epsilon(t) \tag{7.18}$$

where ι^* denotes the indicator function

$$\iota^* = \begin{cases} 1 \text{ if } |\hat{\theta}(t)| = M_p \text{ and } \hat{\theta}^\top(t)\Gamma^x\Omega^\top(t)\epsilon(t) > 0. \\ 0 \text{ otherwise.} \end{cases}$$

The initial weight vector is chosen as $\hat{\theta}(0) = 0$ so that $\hat{\eta}(y, u, \hat{\theta}(0)) = 0$ which corresponds to the case in which the dynamics of the estimator are described only in terms of the known dynamics f.

Remark 7.1. *The signal $\epsilon(t)$ is used in the adaptive law for adjusting the parameter vector $\hat{\theta}$ since it provides a measure of the uncertainty $\Delta\eta(t)$. It is important to note that, from* (7.13)*, a suitable bound for $|\epsilon(t)|$ can be derived and hence, $\epsilon(t)$ can also be used as a residual signal for the task of fault detection. In this case though, $\epsilon(t)$ is generated by using the filter $H_1(s)$ which, although being suitable for the learning task (due to the SPR property), it has a proper transfer function which allows the noise $\xi(t)$ to pass through unaffected. As a result, the bound on $\epsilon(t)$ may be too conservative for practical purposes, and hence it may not be exceeded, thus leading to missed faults. Therefore, the fault detection scheme described in the sequel relies on a general strictly proper transfer function $H(s)$ (through the series of $H_1(s)$ and $H_2(s)$) which possesses better noise dampening characteristics and allows the derivation of tight detection thresholds.*

7.2.3 Residual generation

The residual signal $r(t)$ to be used for fault detection in this chapter is given by

$$r(t) \triangleq H_2(s)\left[\epsilon(t)\right], \tag{7.19}$$

where $H_2(s)$ is given by (7.8) and $\epsilon(t)$ is given by (7.11).

Christodoulos Keliris

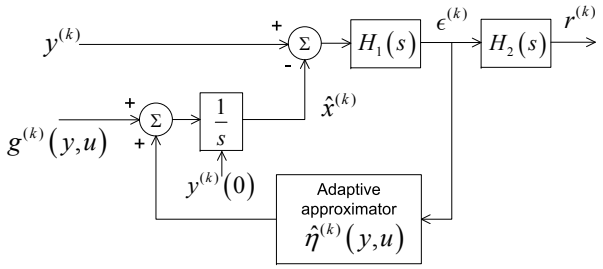


Figure 7.2: Residual Generation

By using (7.11), the residual (7.19) becomes

$$\begin{aligned} r(t) &= H(s)\big[y(t) - \hat{x}(t)\big] \\ &= sH_p(s)\big[x(t) - \hat{x}(t)\big] + H(s)\big[\xi(t)\big] \\ &= H_p(s)\big[\dot{x}(t) - \dot{\hat{x}}(t) + \big(x(0) - \hat{x}(0)\big)\delta(t)\big] + \epsilon_\xi(t) \\ &= H_p(s)\big[\Delta g(t) + \Delta\eta(t) + \beta^x(t - T_0^x)\phi(x(t), u(t))\big] \\ &\quad + H(s)\big[\beta^y(t - T_0^y)\sigma(t)\big] - h_p(t)\xi(0) + \epsilon_\xi(t), \end{aligned} \tag{7.20}$$

where $h_p(t)$ is the impulse response associated with $H_p(s)$.

Under fault-free operation, the residual (7.20) is written as

$$r(t) = H_p(s)\big[\Delta g(t) + \Delta\eta(t)\big] - h_p(t)\xi(0) + \epsilon_\xi(t). \tag{7.21}$$

Figure 7.2 illustrates the residual generation procedure according to equations (7.7), (7.8), (7.10), (7.11) and (7.19).

7.2.4 Detection threshold

In this subsection, the derivation of suitable detection threshold signals, that guarantee no-false alarms, is presented. In order to exploit the benefits from the learning process, it is considered that the learning period is conducted during the time interval $[0, T_L]$ in which it is assumed that no faults occur. In other words, at time T_L the update of the parameter vector $\hat{\theta}(t)$ stops and the fault detection scheme is initiated. The training time T_L is selected by the designer and should be sufficiently large to allow learning of the modeling uncertainty η. The detection decision of a fault in the system is made when $|r^{(k)}(t)| > \bar{r}^{(k)}(t)$ at some time $t > T_L$, for at least one component $k = 1, 2, \ldots, n$, where $\bar{r}^{(k)}(t)$ is the detection threshold, to be designed in the sequel.

For the derivation of the detection threshold we consider that after time T_L, the functional discrepancy between the unknown function η and its approximator $\hat{\eta}$ satisfies the following assumption.

Christodoulos Keliris

***Assumption* 7.5.** *After the end of the learning phase at time T_L, the error between the modeling uncertainty η and the adaptive approximator $\hat{\eta}$, is bounded as follows*

$$|\eta^{(k)}(x(t),u(t)) - \hat{\eta}_L^{(k)}(y(t),u(t))| \leq \bar{\eta}_L^{(k)}(y(t),u(t)),$$

for all $k = 1,2,\ldots,n$, $t > T_L$ and for all $(x,u) \in \mathcal{D}$, where $\bar{\eta}_L^{(k)}$ is a known bounding function and $\hat{\eta}_L^{(k)}(y,u)$ is the functional approximator using the parameter vector obtained at the end of the learning phase, i.e. $\hat{\eta}_L(y,u) \triangleq \Omega(y,u)\hat{\theta}(T_L)$.

The bound $\bar{\eta}_L^{(k)}$ can be computed by applying a set membership technique, as indicated in [74] and the references therein. In addition, the bound $\bar{\eta}_L^{(k)}$ can take a constant value for a given compact region according to the universal approximation property of the RBF networks [24]. Specifically, consider the case of the linearly parametrized network consisting by radial basis functions (RBF) described by $f_{nn}(Z) = W^\top S(Z) = \sum_{i=1}^{Q} w_i s_i(Z)$ where $Z \in \Omega_Z \subset Q^q$ is the input vector, $W = [w_1,\ldots,w_Q]^\top$ is the weight vector, Q is the number of nodes of the network and $S(Z) = [s_1(Z),\ldots,s_Q(Z)]^\top$ is the vector of RBFs. It has been shown in [13, 71] that for any continuous function $f(Z) : \Omega_Z \to R$, where $\Omega_Z \subset R^q$ is a compact set, and for the linear in parameters RBF network $W^\top S(Z)$ with a sufficiently large node number Q, then there exists an ideal constant weight vector W^* such that the approximation error $\epsilon(Z) \triangleq f(Z) - (W^*)^\top S(Z)$ satisfies $|\epsilon(Z)| \leq \epsilon^*$ for all $Z \in \Omega_Z$. Hence, by selecting a sufficiently large set of basis functions, the approximation error can be made arbitrarily small. The main issue, though, is selecting an appropriate time T_L which may pose a challenge for learning the modeling uncertainty. But, the key point for fault detection enhancement is achieved as long as the bound $\bar{\eta}_L^{(k)}(y,u)$ (with the use of learning) is smaller than the bound on $|\eta^{(k)}(x,u)|$ that would be used instead if no learning was used.

Therefore, $\Delta\eta(t)$ from (7.15), for $t > T_L$ becomes $\Delta\eta(t) = \Delta\eta_L(t)$, where

$$\Delta\eta_L(t) \triangleq \eta\big(x(t),u(t)\big) - \hat{\eta}_L\big(y(t),u(t)\big), \tag{7.22}$$

and therefore, the k-th component of the residual (7.21) becomes

$$\begin{aligned} r^{(k)}(t) = H_p(s)\big[\Delta\eta_L^{(k)}(t)\big] + \epsilon_{\Delta g}^{(k)}(t) - h_p(t)\xi^{(k)}(0) \\ + \epsilon_{\xi}^{(k)}(t), \end{aligned} \tag{7.23}$$

where $\epsilon_{\Delta g}^{(k)}(t) \triangleq H_p(s)\big[g^{(k)}\big(x(t),u(t)\big) - g^{(k)}\big(x(t)+\xi(t),u(t)\big)\big]$ is the filtered form of the function discrepancy term that arises due to the measurement noise. Note that $\epsilon_{\Delta g}^{(k)}(t) = H_p(s)\big[\Delta g^{(k)}(t)\big]$ under sensor fault-free operation (due to the use of measurements in $\Delta g(t)$

Christodoulos Keliris

given by (7.14)). For the derivation of the detection threshold, we make the following assumption.

***Assumption* 7.6.** *The filtered function mismatch term $\epsilon_{\Delta g}^{(k)}(t)$ (under sensor fault-free operation) is bounded by a computable positive function $\bar{\epsilon}_{\Delta g}^{(k)}(t)$; i.e., for all $t > 0$*

$$|\epsilon_{\Delta g}^{(k)}(t)| \leq \bar{\epsilon}_{\Delta g}^{(k)}(t), \quad k = 1, 2, \ldots, n.$$

Assumption 7.6 is based on the observation that filtering dampens the error effect of measurement noise present in the function mismatch term $\Delta g(t)$. A suitable selection of the bound $\bar{\epsilon}_{\Delta g}^{(k)}$ can be derived through the use of simulations (i.e. Monte Carlo) by filtering the function mismatch term using the nominal function and the available noise characteristics (note that the measurement noise is assumed to take values in a compact set). Alternatively, one may exploit the Lipschitz property of the function $f^{(k)}(\cdot)$ and derive a suitable bound as $\bar{\epsilon}_{\Delta g}^{(k)} = (|A^{(k)}| + \lambda_{f_k})\bar{H}_p(s)\big[\bar{\xi}_d\big]$, where $A^{(k)}$ indicates the k-th row of A and, $\bar{H}_p(s)$ is a transfer function with an impulse response $\bar{h}_p(t)$ such that $|h_p(t)| \leq \bar{h}_p(t)$ for all $t > 0$. However, this bound is more conservative since the filtering benefits are not exploited.

By using the triangle inequality, the residual (7.23) satisfies

$$|r^{(k)}(t)| \leq |H_p(s)\big[\Delta\eta^{(k)}(t)\big]| + |\epsilon_{\Delta g}^{(k)}(t)| + |h_p(t)\xi^{(k)}(0)| + |\epsilon_{\xi}^{(k)}(t)|.$$

Therefore, by using the well-known result that the impulse response $h_p(t)$ of a strictly proper and asymptotically stable transfer function $H_p(s)$ decays exponentially; i.e., $|h_p(t)| \leq \nu e^{-\zeta t}$ for some $\nu > 0$, $\zeta > 0$, for all $t > 0$ [21], and by using Assumptions 7.5 and 7.6, the detection threshold for $t > T_L$ is given by

$$\bar{r}^{(k)}(t) \triangleq \bar{H}_p(s)\big[\bar{\eta}_L^{(k)}(y(t), u(t))\big] + \bar{\epsilon}_{\Delta g}^{(k)}(t) + |h_p(t)|\bar{\xi}^{(k)} + \bar{\epsilon}_{\xi}^{(k)}(t), \tag{7.24}$$

where $\bar{H}_p(s)$ is a filter with transfer function $\bar{H}_p(s) = \frac{\nu}{s+\zeta}$. Note that, even a conservative bound $\bar{\xi}^{(k)}$ affects the threshold only during the initial transient since $h_p(t)$ is exponentially decaying.

Remark 7.2. *In the previous chapters and some earlier works (i.e. [28, 96]), it was considered that the modeling uncertainty $\eta^{(k)}$ was bounded by an appropriate bounding function $\bar{\eta}^{(k)}$, which could lead to conservative detection thresholds. In this work we exploit the learned modeling uncertainty to derive less conservative thresholds through $\bar{\eta}_L^{(k)}$ which can be smaller than $\bar{\eta}^{(k)}$, in addition to noise attenuation through filtering, and therefore, enhance*

Christodoulos Kallis

fault detectability. For instance, in this works' setting, if we didn't utilize learning or filtering, but instead considered that the modeling uncertainty satisfies $|\eta^{(k)}(x,u)| \leq \bar{\eta}^{(k)}(y,u)$ *for all* $k = 1,\ldots,n$*, then the estimator would simply be* $\dot{\hat{x}} = g(y,u)$*, the residual would directly be given by* $\epsilon(t) = H_1(s)[y(t) - \hat{x}(t)]$ *and a suitable threshold in this case would be* $\bar{\epsilon}^{(k)}(t) = \frac{1}{s+\alpha}\big[\bar{\Delta}g^{(k)}(t) + \bar{\eta}^{(k)}(y(t),u(t)) + \alpha\bar{\xi}^{(k)}(t)\big] + e^{-\alpha t}\bar{\xi}^{(k)}(t) + \bar{\xi}^{(k)}(t)$*, where* $\bar{\Delta}g^{(k)}(t)$ *is a bound on the mismatch function* $\Delta g^{(k)}(t)$*. Therefore, the threshold in this case may be more conservative due to the bounds used (on the modeling uncertainty, the noise and the mismatch function* $\Delta g^{(k)}(t)$*), leading to missed detections. This is further illustrated in the simulation results (Section 7.6).*

7.3 Fault Detectability Analysis

The design and analysis of the fault detection scheme in the previous two sections were based on the derivation of suitable thresholds $\bar{r}^{(k)}(t)$ such that in the absence of any fault, the residual signals $r^{(k)}(t)$ are bounded by $\bar{r}^{(k)}(t)$ for $t > T_L$. In this section, fault detectability conditions for the aforementioned fault detection scheme are derived. The fault detectability analysis provides an implicit characterization, in non-closed form, of the class of faults detectable by the proposed scheme.

Theorem 7.1. *Consider the nonlinear system (7.1), (7.2) with the adaptive approximation structure described in* (7.16)*,* (7.17)*,* (7.7)*,* (7.10)*,* (7.11) *and the fault detection scheme described in* (7.8)*,* (7.19)*,* (7.24)*.*

a) Process fault detectability. A process fault occurring at time $T_0^x > T_L$ *is detectable at time* $T_d > T_0^x$ *if the following condition is satisfied for some* $k = 1,2,\ldots,n$*:*

$$|H_p(s)\big[\beta^x(T_d - T_0^x)\phi^{(k)}(x(T_d),u(T_d))\big]| > 2\bar{r}^{(k)}(T_d). \tag{7.25}$$

b) Sensor fault detectability. A sensor fault occurring at time $T_0^y > T_L$ *is detectable at time* $T_d > T_0^y$ *if the following condition is satisfied for some* $k = 1,2,\ldots,n$*:*

$$\begin{aligned}|H(s)\big[\sigma^{(k)}(T_d)\big] + H_p(s)\big[g^{(k)}(x(T_d)+\xi(T_d),u(T_d))\\ - g^{(k)}(x(T_d)+\xi(T_d)+\sigma(T_d),u(T_d))\big]| > 2\bar{r}^{(k)}(T_d).\end{aligned} \tag{7.26}$$

Proof. **(a)** In the presence of a process fault that occurs at some time $T_0^x > T_L$, the residual from (7.20) for $t > T_0^x$ becomes

$$r(t) = H_p(s)\big[\Delta g(t) + \Delta\eta_L(t) + \beta^x(t - T_0^x)\phi(x(t),u(t))\big] - h_p(t)\xi(0) + \epsilon_\xi(t), \tag{7.27}$$

Christodoulos Kelinis

By using the triangle inequality, the k-th component of (7.27) satisfies

$$\begin{aligned}
|r^{(k)}(t)| &\geq -|H_p(s)\big[\Delta\eta_L^{(k)}(t)\big]| - |\epsilon_{\Delta g}^{(k)}(t)| - |h_p(t)||\xi^{(k)}(0)| \\
&\quad - |\epsilon_{\xi}^{(k)}(t)| + |H_p(s)\big[\beta^x(t-T_0^x)\phi^{(k)}(x(t),u(t))\big]| \\
&\geq -\bar{H}_p(s)\big[\bar{\eta}_L^{(k)}(y(t),u(t))\big] - \bar{\epsilon}_{\Delta g}^{(k)}(t) - |h_p(t)|\bar{\xi}^{(k)} \\
&\quad - \bar{\epsilon}_{\xi}^{(k)}(t) + |H_p(s)\big[\beta^x(t-T_0^x)\phi^{(k)}(x(t),u(t))\big]| \\
&\geq -\bar{r}^{(k)}(t) + |H_p(s)\big[\beta^x(t-T_0^x)\phi^{(k)}(x(t),u(t))\big]|.
\end{aligned}$$

Therefore, for fault detection the inequality $|r^{(k)}(t)| > \bar{r}^{(k)}(t)$ must hold at some time $t = T_d$ for some $k = 1, 2, \ldots, n$, so the final fault detectability condition in (7.25) is obtained.

(b) In the presence of a sensor fault that occurs at some time $T_0^y > T_L$, the residual from (7.20) by using (7.14) and, by adding and subtracting $g(x(t)+\xi(t), u(t))$ for $t > T_0^y$ becomes

$$\begin{aligned}
r(t) = H_p(s)\big[&g(x(t),u(t)) - g(x(t)+\xi(t),u(t)) \\
&+ g(x(t)+\xi(t),u(t)) - g(x(t)+\xi(t)+\sigma(t),u(t))\big] \\
&+ H_p(s)\big[\Delta\eta_L(t)\big] - h_p(t)\xi(0) + \epsilon_\xi(t) + H(s)\big[\sigma(t)\big]. \qquad (7.28)
\end{aligned}$$

By using the triangle inequality and Assumptions 7.5, 7.6, the k-th component of (7.28) satisfies

$$\begin{aligned}
|r^{(k)}(t)| &\geq \big|H(s)\big[\sigma^{(k)}(t)\big] + H_p(s)\big[g^{(k)}(x(t)+\xi(t),u(t)) - g^{(k)}(x(t)+\xi(t)+\sigma(t),u(t))\big]\big| \\
&\quad - |\epsilon_{\Delta g}^{(k)}(t)| - |H_p(s)\big[\Delta\eta_L^{(k)}(t)\big]| - |h_p(t)||\xi^{(k)}(0)| - |\epsilon_{\xi}^{(k)}(t)| \\
&\geq \big|H(s)\big[\sigma^{(k)}(t)\big] + H_p(s)\big[g^{(k)}(x(t)+\xi(t),u(t)) - g^{(k)}(x(t)+\xi(t)+\sigma(t),u(t))\big]\big| \\
&\quad - \bar{\epsilon}_{\Delta g}^{(k)}(t) - \bar{H}_p(s)\big[\bar{\eta}_L^{(k)}(y(t),u(t))\big] - |h_p(t)|\bar{\xi}^{(k)} - \bar{\epsilon}_{\xi}^{(k)}(t) \\
&\geq \big|H(s)\big[\sigma^{(k)}(t)\big] + H_p(s)\big[g^{(k)}(x(t)+\xi(t),u(t)) - g^{(k)}(x(t)+\xi(t)+\sigma(t),u(t))\big]\big| \\
&\quad - \bar{r}^{(k)}(t).
\end{aligned}$$

Therefore, for fault detection the inequality $|r^{(k)}(t)| > \bar{r}^{(k)}(t)$ must hold at some time $t = T_d$ for some $k = 1, 2, \ldots, n$, so the final fault detectability condition in (7.26) is obtained. □

The above theorem provides sufficient conditions for the implicit characterization of a class of faults that can be detected by the proposed fault detection scheme. As seen from the detectability conditions, the detection of the fault depends on the filtered fault functions. This demonstrates the crucial importance of the filter selection, which allows the derivation of tight detection thresholds. In addition, note that the detectability conditions, which are obtained under "worst case" conditions, are only sufficient (but not necessary) and hence, the classes of the detectable faults may be significantly larger.

Christodoulos Keliris

7.4 Fault Type Identification

In this section we proceed to the task of identification of the fault type and the estimation of the magnitude of the fault that has occurred. When a fault is detected at time T_d, we initiate two distinct estimation models (FIEs) in order to identify the potential fault that has occurred. The first estimation model considers that a process fault has occurred and relies on the same logic for learning the modeling uncertainty by resuming learning to allow learning of the process fault as well. The second estimation model considers the occurrence of a sensor fault and makes use of learning to estimate adaptively the potential sensor fault(s) magnitude that occurred. The fault type identification methodology relies on an exclusion-based logic, i.e. suitable threshold signals in both estimation models are obtained and when they are exceeded, the particular fault type is excluded and hence, the occurrence of the the other fault type is guaranteed. Furthermore, the use of learning in both estimation models provides an estimation of the fault that has occurred once the fault type identification has been concluded.

7.4.1 Process Fault

In this part we consider the case of a process fault. When a fault is detected at time T_d, we use the same logic that was used to learn the modeling uncertainty, with the only difference that this time the task is to learn the combined effect of the modeling uncertainty coupled with the potential process fault function. Essentially, we use the same estimation model as before indicated by (7.10), (7.11), (7.19) and resume learning. For clarity, we rewrite the equations as follows

$$\dot{\hat{x}}_p(t) = g(y(t), u(t)) + \hat{\eta}(y(t), u(t), \hat{\theta}(t)) \tag{7.29}$$

$$\epsilon_p(t) \triangleq H_1(s)\big[y(t) - \hat{x}_p(t)\big] \tag{7.30}$$

$$r_p(t) \triangleq H_2(s)\left[\epsilon_p(t)\right], \tag{7.31}$$

with initial conditions $\hat{x}_p(T_d) = y(T_d)$, $\hat{\theta}(T_d) = \hat{\theta}(T_L)$ (resume learning). As before, we allow a training time $T_{L,2}$ to learn the combined effect of the modeling uncertainty and the potential process fault., i.e. the training is conducted during the interval $[T_d, T_d + T_{L,2}]$, and make the following assumption, similarly to Assumption 7.5.

***Assumption* 7.7.** *In the case of a process fault, after the end of the learning phase at time* $T_d + T_{L,2}$*, the error between the combined modeling uncertainty and fault function* $\eta(\cdot) + \beta^x(\cdot)\phi(\cdot)$

Christodoulos Keliris

and the adaptive approximator $\hat{\eta}$ i.e. $\Delta\eta_{L,2}(t) \triangleq \eta(x(t),u(t)) + \beta^x(t-T_0^x)\phi(x(t),u(t)) - \hat{\eta}_{L,2}(y(t),u(t))$*, is bounded as follows*

$$|\Delta\eta_{L,2}^{(k)}(t)| \leq \bar{\eta}_{L,2}^{(k)}(y(t),u(t)),$$

for all $k = 1,2,\ldots,n$, $t > T_d + T_{L,2}$ and for all $(x,u) \in \mathcal{D}$, where $\bar{\eta}_{L,2}^{(k)}$ is a known bounding function and $\hat{\eta}_{L,2}^{(k)}(y,u)$ is the functional approximator using the parameter vector obtained at the end of the second stage of the learning phase; i.e., $\hat{\eta}_{L,2}(y,u) \triangleq \Omega(y,u)\hat{\theta}(T_d + T_{L,2})$.

Therefore, in the case of a process fault, by using (7.29), (7.30), (7.31), r_p satisfies the following equation for $t > T_d + T_{L,2}$ (derived similarly to (7.21)):

$$r_p(t) = H_p(s)\big[\Delta\eta_{L,2}(t)\big] + H_p(s)\big[\Delta g(t)\big] - h_p(t)\xi(T_d) + \epsilon_\xi(t). \tag{7.32}$$

Hence, by using the triangle inequality and by using Assumption 7.6 (note that it is valid under sensor fault-free operation) and Assumption 7.7, the k-th component of r_p from (7.32) satisfies:

$$|r_p^{(k)}(t)| \leq |H_p(s)\big[\Delta\eta_{L,2}^{(k)}(t)\big]| + |\epsilon_{\Delta g}^{(k)}(t)| + |h_p(t)||\xi^{(k)}(T_d)| + |\epsilon_\xi^{(k)}(t)|$$

and hence, $|r_p^{(k)}(t)| \leq \bar{r}_p^{(k)}(t)$ where $\bar{r}_p^{(k)}(t)$ is the process isolation threshold given by

$$\bar{r}_p^{(k)}(t) \triangleq \bar{H}_p(s)\big[\bar{\eta}_{L,2}^{(k)}(y(t),u(t))\big] + \bar{\epsilon}_{\Delta g}^{(k)}(t) + |h_p(t)|\bar{\xi}^{(k)} + \bar{\epsilon}_\xi^{(k)}(t). \tag{7.33}$$

Fault Identification Logic: When a process fault occurs and it is detected at time T_d, then the residual $r_p(t)$ given by (7.29), (7.30), (7.31), satisfies $|r_p^{(k)}(t)| \leq \bar{r}_p^{(k)}(t)$ for all $k = 1,\ldots,n$ and for all $t > T_d + T_{L,2}$. Hence, if at some time $t > T_d + T_{L,2}$, we have $|r_p^{(k)}(t)| > \bar{r}_p^{(k)}(t)$ for some $k = 1,\ldots,n$ then the possibility of a process fault is excluded and it is guaranteed that a sensor fault has occurred (single or multiple sensor fault(s)). The estimation of the sensor fault is given by $\hat{\sigma}(t)$ (to be designed in the sequel) from which the faulty sensors can be identified as the components that deviate significantly from zero.

7.4.2 Sensor Fault

In this part we consider the occurrence of a sensor fault. When a fault is detected at time T_d, the following estimation model is initiated to learn the potential sensor fault that occurred:

$$\dot{\hat{x}}_s(t) = A\hat{x}_s(t) + f(\hat{x}_s(t),u(t)) + \hat{\eta}_L(y(t),u(t)) + \Lambda\epsilon_s^y(t) \tag{7.34}$$

$$\hat{y}_s(t) = \hat{x}_s(t) + \hat{\sigma}(t) \tag{7.35}$$

$$\dot{\hat{\sigma}}(t) = \mathcal{P}_\sigma\big(\Gamma^y\epsilon_s^y(t)\big), \tag{7.36}$$

Christodoulos Keliris

where $\epsilon_s^y(t)$ denotes the output estimation error $\epsilon_s^y(t) \triangleq y(t) - \hat{y}_s(t)$, Λ is a design matrix that is selected so that $A_0 \triangleq A - \Lambda$ is Hurwitz, $\Gamma^y \in \mathbb{R}^{n\times n}$ is a symmetric and positive definite learning rate matrix and $\hat{\sigma}(t) \in \mathbb{R}^n$ denotes the sensor fault estimation. Finally, the initial conditions are $\hat{x}_s(T_d) = 0$ and $\hat{\sigma}(T_d) = 0$. In the following, we denote the state estimation error as $\epsilon_s^x(t) \triangleq x(t) - \hat{x}_s(t)$. The projection operator $\mathcal{P}_\sigma$ restricts the estimation vector $\hat{\sigma}(t)$ in a predefined and convex region $\Theta_\sigma \in \mathbb{R}^n$ in order to guarantee stability of the learning algorithm in the presence of noise and modeling uncertainty. In this work, Θ_σ is considered to be a zero-origin hypersphere of radius M_σ, and hence the adaptive law for $\dot{\hat{\sigma}}$ can be expressed as [83, 95]

$$\dot{\hat{\sigma}}(t) = \Gamma^y \epsilon_s^y(t) - \chi^* \Gamma^y \frac{\hat{\sigma}(t)\hat{\sigma}^\top(t)}{\hat{\sigma}^\top(t)\Gamma^y\hat{\sigma}(t)} \Gamma^y \epsilon_s^y(t) \tag{7.37}$$

where χ^* denotes the indicator function

$$\chi^* = \begin{cases} 1 \text{ if } |\hat{\sigma}(t)| = M_\sigma \text{ and } \hat{\sigma}^\top(t)\Gamma^y\epsilon_s^y(t) > 0. \\ 0 \text{ otherwise.} \end{cases} \tag{7.38}$$

The following main result summarizes the properties of the aforementioned adaptive estimation scheme.

Theorem 7.2. *In the case of a sensor fault $\sigma(t)$ that occurs at time T_0^y and is detected at time T_d, the adaptive nonlinear estimation scheme described by* (7.34)*,* (7.35)*,* (7.36) *under Assumptions 7.1, 7.2, 7.3 and 7.5, with the positive constants μ, ρ selected such that $|e^{A_0 t}| \leq \mu e^{-\rho t}$ for all $t > 0$ and $\rho - \mu\lambda_f > 0$, guarantees that:*

(a) The state estimation error $\epsilon_s^x(t)$, the output estimation error $\epsilon_s^y(t)$ and the sensor fault estimation vector $\hat{\sigma}(t)$ are uniformly bounded, i.e. $\epsilon_s^x(t)$, $\epsilon_s^y(t)$, $\hat{\sigma}(t) \in L_\infty$. Specifically, a bound on $\epsilon_s^x(t)$ such that $|\epsilon_s^x(t)| \leq \bar{\epsilon}_s^x(t)$ for all $t > T_d$ is given by (7.49) *and a bound for $\epsilon_s^y(t)$ such that $|\epsilon_s^y(t)| \leq \bar{\epsilon}_s^y(t)$ for all $t > T_d$ is given by* (7.51)*.*

(b) There exist a positive constant q and a bounded function $v(t)$ such that, for all finite $t > T_d$ the output estimation error $\epsilon_s^y(t)$ satisfies:

$$\int_{T_d}^{t} |\epsilon_s^y(\tau)|^2 d\tau \leq q + 4\int_{T_d}^{t} |v(\tau)|^2 d\tau. \tag{7.39}$$

(c) In the absence of measurement noise (i.e. $\xi(t) = 0$), if the bounded function $v(t)$ is square integrable (i.e. $v(t) \in L_2$), then $\lim_{t\to\infty} \epsilon_s^y(t) = 0$.

Christodoulos Kellinis

Proof. **(a)** When a sensor fault occurs, by using (7.1) (note that $\phi(\cdot) = 0$) and (7.34) the state estimation error $\epsilon_s^x(t) = x(t) - \hat{x}_s(t)$ satisfies

$$\begin{aligned}\dot{\epsilon}_s^x(t) &= \dot{x}(t) - \dot{\hat{x}}_s(t) \\ &= Ax(t) + f(x(t), u(t)) + \eta(x(t), u(t)) - A\hat{x}_s(t) \\ &\quad - f(\hat{x}_s(t), u(t)) - \hat{\eta}_L(y(t), u(t)) - \Lambda(y(t) - \hat{y}_s(t)) \qquad (7.40) \\ &= A\epsilon_s^x(t) + \Delta f_s(t) + \Delta\eta_L(t) - \Lambda\big(x(t) + \xi(t) + \sigma(t) - \hat{x}_s(t) - \hat{\sigma}(t)\big), \\ &= A_0\epsilon_s^x(t) + \Lambda\tilde{\sigma}(t) - \Lambda\xi(t) + \Delta f_s(t) + \Delta\eta_L(t), \qquad (7.41)\end{aligned}$$

where

$$\Delta f_s(t) \triangleq f(x(t), u(t)) - f(\hat{x}_s(t), u(t)) \qquad (7.42)$$

$$\tilde{\sigma}(t) \triangleq \hat{\sigma}(t) - \sigma(t), \qquad (7.43)$$

and $\Delta\eta_L(t)$ is given by (7.22) (note that a fault is considered to occur after the end of training at time T_L). By solving (7.41) we obtain

$$\begin{aligned}\epsilon_s^x(t) &= e^{A_0(t-T_d)}\epsilon_s^x(T_d) \\ &\quad + \int_{T_d}^{t} e^{A_0(t-\tau)}\big[\Lambda\tilde{\sigma}(\tau) - \Lambda\xi(\tau) + \Delta f_s(\tau) + \Delta\eta_L(\tau)\big]\, d\tau. \qquad (7.44)\end{aligned}$$

Using the triangle inequality, (7.44) satisfies

$$\begin{aligned}|\epsilon_s^x(t)| &\leq |e^{A_0(t-T_d)}||\epsilon_s^x(T_d)| \\ &\quad + \int_{T_d}^{t} |e^{A_0(t-\tau)}|\big[|\Lambda||\tilde{\sigma}(\tau)| + |\Lambda||\xi(\tau)| + |\Delta f_s(\tau)| + |\Delta\eta_L(\tau)|\big]\, d\tau. \qquad (7.45)\end{aligned}$$

Since A_0 is Hurwitz, there exist positive constants μ, ρ such that $|e^{A_0 t}| \leq \mu e^{-\rho t}$ for all $t > 0$ [43]. Moreover, by using the initial conditions $\hat{x}_s(T_d) = 0$, we obtain $|\epsilon_s^x(T_d)| = |x(T_d) - \hat{x}_s(T_d)| = |x(T_d)| \leq \bar{x}_d$, where $\bar{x}_d$ is a suitable bound for the region of the state operation. By using Assumption 7.2, $|\Delta f_s(t)| \leq \lambda_f |x(t) - \hat{x}_s(t)| = \lambda_f |\epsilon_s^x(t)|$ and, by using Assumption 7.5 we have that $|\Delta\eta_L(t)| \leq |\bar{\eta}_L(y(t), u(t))|$. Regarding the term $\tilde{\sigma}(t) = \hat{\sigma}(t) - \sigma(t)$, it is generally unknown but, since $\hat{\sigma}(t)$ belongs to the known compact set Θ_σ, we have that $|\hat{\sigma}(t) - \sigma(t)| \leq \kappa(t)$ for a suitable $\kappa(t)$ depending on the geometric properties of the set Θ_σ (see [95, 96]). More specifically, in our case where the compact set Θ_σ is an origin-based hypersphere of radius M_σ, the function $\kappa(t)$ is given by $\kappa(t) = M_\sigma + |\hat{\sigma}(t)|$. Hence, (7.45) becomes

$$|\epsilon_s^x(t)| \leq \mu e^{-\rho(t-T_d)}\bar{x}_d + \int_{T_d}^{t} \mu e^{-\rho(t-\tau)}\big[|\Lambda|\kappa(\tau) + |\Lambda|\bar{\xi}_d + \lambda_f|\epsilon_s^x(\tau)| + |\bar{\eta}_L(y(\tau), u(\tau))|\big]\, d\tau, \qquad (7.46)$$

Christodoulos Keliris

which can be written as

$$|\epsilon_s^x(t)| \leq E_0(t) + \mu\lambda_f e^{-\rho t} \int_{T_d}^{t} e^{\rho\tau} |\epsilon_s^x(\tau)|\, d\tau, \tag{7.47}$$

where

$$E_0(t) \triangleq \mu e^{-\rho(t-T_d)} \bar{x}_d + \int_{T_d}^{t} \mu e^{-\rho(t-\tau)} \big[|\Lambda| \kappa(\tau) + |\Lambda| \bar{\xi}_d + |\bar{\eta}_L(y(\tau), u(\tau))| \big] d\,\tau. \tag{7.48}$$

By applying the Bellman-Gronwall lemma [43] to (7.47), it results in $|\epsilon_s^x(t)| \leq \bar{\epsilon}_s^x(t)$ where

$$\bar{\epsilon}_s^x(t) \triangleq E_0(t) + \mu\lambda_f \int_{T_d}^{t} E_0(w) e^{-(\rho - \mu\lambda_f)(t-w)}\, dw, \tag{7.49}$$

where the positive constants μ, ρ must be selected such that $\rho - \mu\lambda_f > 0$ to guarantee that $\epsilon_s^x(t)$ remains bounded, hence $\epsilon_s^x(t) \in L_\infty$. Moreover, since $|\tilde{\sigma}(t)| = |\hat{\sigma}(t) - \sigma(t)| \leq \kappa(t)$ due to the projection operator, we deduce that $\tilde{\sigma}(t) \in L_\infty$, $\hat{\sigma}(t) \in L_\infty$. Finally, by using (7.2), (7.35), $\epsilon_s^y(t) = y(t) - \hat{y}_s(t)$ becomes

$$\epsilon_s^y(t) = \epsilon_s^x(t) + \xi(t) - \tilde{\sigma}(t). \tag{7.50}$$

Using the triangle inequality, (7.50) satisfies,

$$\begin{aligned} |\epsilon_s^y(t)| &\leq |\epsilon_s^x(t)| + |\xi(t)| + |\tilde{\sigma}(t)| \\ &\leq \bar{\epsilon}_s^y(t) \triangleq \bar{\epsilon}_s^x(t) + \bar{\xi}_d + \kappa(t), \end{aligned} \tag{7.51}$$

and hence, $\epsilon_s^y(t)$ is bounded, i.e. $\epsilon_s^y(t) \in L_\infty$.

(b) From (7.44), $\epsilon_s^x(t)$ can be written as

$$\epsilon_s^x(t) = \upsilon_1(t) + \upsilon_2(t), \tag{7.52}$$

where $\upsilon_1(t)$, $\upsilon_2(t)$ are the solutions of

$$\begin{aligned} \dot{\upsilon}_1(t) &= -\Lambda\upsilon_1(t) + \Lambda\tilde{\sigma}(t) - \Lambda\xi(t) + \Delta f_s(t) + \Delta\eta_L(t) \\ \dot{\upsilon}_2(t) &= -\Lambda\upsilon_2(t) \end{aligned}$$

with initial conditions $\upsilon_1(T_d) = 0$ and $\upsilon_2(T_d) = \epsilon_s^x(T_d) = x(T_d)$ (note that $\hat{x}_s(T_d) = 0$).

By using (7.52), (7.50) becomes

$$\epsilon_s^y(t) = \upsilon_1(t) + \upsilon_2(t) + \xi(t) - \tilde{\sigma}(t). \tag{7.53}$$

Let the Lyapunov function candidate be $V(t) = \frac{1}{2}\tilde{\sigma}^\top(t)(\Gamma^y)^{-1}\tilde{\sigma}(t) + \int_t^\infty |\upsilon_2(\tau)|^2\, d\tau$. The time derivative of $V(t)$ is given by

$$\begin{aligned} \dot{V}(t) &= \tilde{\sigma}^\top(t)(\Gamma^y)^{-1}\dot{\tilde{\sigma}}(t) - |\upsilon_2(t)|^2 \\ &= \tilde{\sigma}^\top(t)(\Gamma^y)^{-1}\big(\dot{\hat{\sigma}}(t) - \dot{\sigma}(t)\big) - |\upsilon_2(t)|^2, \end{aligned}$$

Christodoulos Kaliris

and by using (7.37), it becomes

$$\dot{V}(t) = \tilde{\sigma}^\top(t)\epsilon_s^y(t) - \tilde{\sigma}^\top(t)\chi^* \frac{\hat{\sigma}(t)\hat{\sigma}^\top(t)}{\hat{\sigma}^\top(t)\Gamma^y\hat{\sigma}(t)}\Gamma^y\epsilon_s^y(t) - \tilde{\sigma}^\top(t)(\Gamma^y)^{-1}\dot{\sigma}(t) - |\upsilon_2(t)|^2. \tag{7.54}$$

Following the same logic as in [83], it can be shown that the second term in (7.54) satisfies

$$\tilde{\sigma}^\top(t)\chi^* \frac{\hat{\sigma}(t)\hat{\sigma}^\top(t)}{\hat{\sigma}^\top(t)\Gamma^y\hat{\sigma}(t)}\Gamma^y\epsilon_s^y(t) \geq 0$$

and hence, (7.54) can be written as

$$\dot{V}(t) \leq \tilde{\sigma}^\top(t)\epsilon_s^y(t) - \tilde{\sigma}^\top(t)(\Gamma^y)^{-1}\dot{\sigma}(t) - |\upsilon_2(t)|^2. \tag{7.55}$$

By using $\tilde{\sigma}(t) = -\epsilon_s^y(t) + \upsilon_1(t) + \upsilon_2(t) + \xi(t)$ (from (7.53)) and by completing the squares, (7.55) becomes

$$\begin{aligned}
\dot{V}(t) &\leq (\epsilon_s^y(t))^\top(-\epsilon_s^y(t) + \upsilon_1(t) + \upsilon_2(t) + \xi(t)) - \tilde{\sigma}^\top(t)(\Gamma^y)^{-1}\dot{\sigma}(t) - |\upsilon_2(t)|^2 \\
&\leq -|\epsilon_s^y(t)|^2 + |\epsilon_s^y(t)|\bar{\xi}_d + |\epsilon_s^y(t)||\upsilon_1(t)| + |\epsilon_s^y(t)||\upsilon_2(t)| \\
&\quad - |\upsilon_2(t)|^2 + |\tilde{\sigma}(t)||(\Gamma^y)^{-1}||\dot{\sigma}(t)| \\
&\leq -\frac{|\epsilon_s^y(t)|^2}{4} + \bar{\xi}_d^2 + |\upsilon_1(t)|^2 + |\tilde{\sigma}(t)||(\Gamma^y)^{-1}||\dot{\sigma}(t)|,
\end{aligned} \tag{7.56}$$

Let $v(t) \triangleq \big(|\upsilon_1(t)|^2 + \bar{\xi}_d^2 + |\tilde{\sigma}(t)||(\Gamma^y)^{-1}||\dot{\sigma}(t)|\big)^{\frac{1}{2}}$ (note that v is bounded). Then, by integrating (7.56) from time T_d to t we obtain

$$\begin{aligned}
\int_{T_d}^t |\epsilon_s^y(\tau)|^2\, d\tau &\leq 4\big(V(T_d) - V(t)\big) + 4\int_{T_d}^t \big(|\upsilon_1(\tau)|^2 + \bar{\xi}_d^2 + |\tilde{\sigma}(\tau)||(\Gamma^y)^{-1}||\dot{\sigma}(\tau)|\big)\, d\tau \\
&\leq q + 4\int_{T_d}^t |v(\tau)|^2\, d\tau,
\end{aligned} \tag{7.57}$$

where $q \triangleq \sup_{t\geq T_d}\big[4\big(V(T_d) - V(t)\big]$ is a positive constant (since $V(t)$ is uniformly bounded).

(c) Since $\epsilon_s^y(t) \in L_\infty$ as it was shown in part (a) of the proof, then from (7.36) we deduce that $\dot{\hat{\sigma}}(t) \in L_\infty$. Moreover, all terms in the right hand side of (7.41) were shown to be bounded and therefore, $\dot{\epsilon}_s^x(t) \in L_\infty$. Moreover, in the absence of noise (i.e. $\xi(t) = 0$), from (7.50) we have that $\dot{\epsilon}_s^y(t) = \dot{\epsilon}_s^x(t) + \dot{\sigma}(t) - \dot{\hat{\sigma}}(t)$ and, since $\dot{\epsilon}_s^x(t) \in L_\infty$, $\dot{\sigma}(t) \in L_\infty$ (from Assumption 7.3), $\dot{\hat{\sigma}}(t) \in L_\infty$, we obtain that $\dot{\epsilon}_s^y(t) \in L_\infty$.

Therefore, if the bounded function $v(t)$ is square integrable, i.e. $v(t) \in L_2$, then (7.57) can be valid for $t \to \infty$, leading to $\epsilon_s^y(t) \in L_2$. Since $\epsilon_s^y(t) \in L_\infty$, $\dot{\epsilon}_s^y(t) \in L_\infty$, $\epsilon_s^y(t) \in L_2$ then according to Barbalat's Lemma $\lim_{t\to\infty} \epsilon_s^y(t) = 0$. □

Christodoulos Keliris

Now, we exploit the findings of Theorem 7.2 to derive suitable bounds to aid in the determination of the fault type. More specifically, we proceed to derive a threshold for the case of a sensor fault, so that when it is exceeded then it is guaranteed that a process fault has occurred. Two such residuals are given. The first one is the output estimation error $\epsilon_s^y(t) = y(t) - \hat{y}_s(t)$ since it is bounded by $\bar{\epsilon}_s^y(t)$ given by (7.51) for all $t > T_d$ in the case of a sensor fault. The second residual $r_s(t)$ exploits both the use of filtering and some signals from the designed adaptive scheme, to cancel out some terms in the residual and, allow a potentially tighter threshold. Therefore, the residual $r_s(t)$ in this case is obtained by

$$r_s(t) \triangleq H(s)\big[\epsilon_s^y(t) + w(t)\big] \tag{7.58}$$

$$\dot{w}(t) = \Lambda(y(t) - \hat{y}_s(t)) + \dot{\hat{\sigma}}(t), \tag{7.59}$$

with initial condition $w(T_d)$=0. For simplicity, in the sequel, we omit the initial condition terms (now at time T_d) which are exponentially decaying to zero, since they are multiplied with $h_p(t)$, and do not affect substantially the threshold derivation. Therefore, by using (7.50), (7.58) can be written

$$\begin{aligned} r_s(t) &= H(s)\big[\epsilon_s^x(t) + \xi(t) - \tilde{\sigma}(t) + w(t)\big] \\ &= H_p(s)\big[\dot{\epsilon}_s^x(t) - \dot{\tilde{\sigma}}(t) + \dot{w}(t)\big] + \epsilon_\xi(t), \end{aligned} \tag{7.60}$$

and by using (7.40) and (7.59), (7.60) becomes

$$\begin{aligned} r_s(t) = H_p(s)\big[&A\epsilon_s^x(t) + \Delta f_s(t) + \Delta\eta_L(t) - \Lambda(y(t) - \hat{y}_s(t)) \\ &- \big(\dot{\hat{\sigma}}(t) - \dot{\sigma}(t)\big) + \Lambda(y(t) - \hat{y}_s(t)) + \dot{\hat{\sigma}}(t)\big] + \epsilon_\xi(t) \\ = H_p(s)\big[&A\epsilon_s^x(t) + \Delta f_s(t) + \Delta\eta_L(t) + \dot{\sigma}(t)\big] + \epsilon_\xi(t). \end{aligned} \tag{7.61}$$

Thus, by using the triangle inequality, the k-th component of (7.61) satisfies

$$|r_s^{(k)}(t)| \leq \bar{H}_p(s)\big[|A^{(k)}||\epsilon_s^x(t)| + |\Delta f_s^{(k)}(t)| + |\Delta\eta_L^{(k)}(t)| + |\dot{\sigma}^{(k)}(t)|\big] + |\epsilon_\xi^{(k)}(t)|. \tag{7.62}$$

Using Assumption 7.2, 7.3 and 7.5, a suitable threshold $\bar{r}_s^{(k)}(t)$, such that $|r_s^{(k)}(t)| \leq \bar{r}_s^{(k)}(t)$ for all $t > T_d$ and for all $k = 1, \ldots, n$ is obtained as

$$\bar{r}_s^{(k)}(t) \triangleq \bar{H}_p(s)\big[(|A^{(k)}| + \lambda_{f_k})\bar{\epsilon}_s^x(t) + \bar{\eta}_L^{(k)}(y(t), u(t)) + \psi\big] + \bar{\epsilon}_\xi^{(k)}(t). \tag{7.63}$$

Fault Identification Logic: When a sensor fault occurs and is detected at time T_d, then $|\epsilon_s^y(t)| \leq \bar{\epsilon}_s^y(t)$ and $|r_s^{(k)}(t)| \leq \bar{r}_s^{(k)}(t)$ for all $k = 1, \ldots, n$ and for all $t > T_d$. Therefore, if at some time $t > T_d$ we have $|\epsilon_s^y(t)| > \bar{\epsilon}_s^y(t)$ or $|r_s^{(k)}(t)| > \bar{r}_s^{(k)}(t)$ for some $k = 1, \ldots, n$, then the possibility of a sensor fault is excluded and it is guaranteed that a process fault has occurred. The estimation of the process fault (including the modeling uncertainty) is given by $\hat{\eta}_{L,2}$ for $t > T_d + T_{L_2}$.

Christodoulos Keliris

7.5 Fault Type Identification Analysis

In this section, fault type identification conditions for the aforementioned fault diagnosis scheme are derived. This analysis, similarly to the fault detectability analysis given in Section 7.3, constitutes a theoretical result that characterizes quantitatively and in non-closed form the classes of faults that can be identified (process or sensor) by the proposed scheme.

Theorem 7.3. *Consider the nonlinear system (7.1), (7.2) with the adaptive approximation structure described in* (7.16)*,* (7.17)*,* (7.7)*,* (7.10)*,* (7.11) *along with the two estimation models: for process faults given by* (7.29)-(7.31) *and for sensor faults given by* (7.34)-(7.36) *and their corresponding isolation thresholds given by* (7.33) *($\bar{r}_p(t)$) and* (7.63) *($\bar{r}_s(t)$).*

a) Process fault identification. A process fault type is identified if one of the following conditions (7.64) *or* (7.66) *is satisfied for some $k = 1, 2, \ldots, n$ at some time $t > T_d$ (the possibility of a sensor fault is excluded):*

$$|\epsilon^x_{s,2}(t)| > \bar{\epsilon}^y_s(t) + \bar{\xi}_d + |\hat{\sigma}(t)| \tag{7.64}$$

where

$$\epsilon^x_{s,2}(t) \triangleq e^{A_0(t-T_d)}x(T_d) + \int_{T_d}^{t} e^{A_0(t-\tau)}\big[\Lambda\hat{\sigma}(\tau) - \Lambda\xi(\tau) + \Delta f_s(\tau) + \Delta\eta_L(\tau) + \beta^x(\tau - T_0^x)\phi(x(\tau), u(\tau))\big]\, d\tau, \tag{7.65}$$

or

$$|H_p(s)\big[\beta^x(t - T_0^x)\phi^{(k)}(x(t), u(t))\big]| > \bar{r}_s^{(k)}(t) + |H_p(s)\big[A^{(k)}\epsilon^x_s(t) + \Delta f_s^{(k)}(t) + \Delta\eta_L^{(k)}(t)\big] + \epsilon_\xi^{(k)}(t)|. \tag{7.66}$$

b) Sensor fault identification. A sensor fault type is identified if the following condition is satisfied for some $k = 1, 2, \ldots, n$ at some time $t > T_d + T_{L,2}$ (the possibility of a process fault is excluded):

$$|H(s)\big[\sigma^{(k)}(t)\big] + H_p(s)\big[g^{(k)}(x(t) + \xi(t), u(t)) - g^{(k)}(x(t) + \xi(t) + \sigma(t), u(t))\big]| > 2\bar{r}_p^{(k)}(t). \tag{7.67}$$

Proof. **(a)** When a fault is detected at time T_d, one of the estimators initiated for identifying the fault type is the one for sensor faults given in Section 7.4.2, which satisfies $|\epsilon^y_s(t)| \leq \bar{\epsilon}^y_s(t)$ and $|r_s^{(k)}(t)| \leq \bar{r}_s^{(k)}(t)$ for all $k = 1, \ldots, n$ and for all $t > T_d$, if a sensor fault has

Christodoulos Keliris

occurred. So, when a process fault occurs, and the fault is detected, the state estimation error $\epsilon_s^x(t) = x(t) - \hat{x}_s(t)$ (now indicated as $\epsilon_{s,2}^x(t)$) is given by (7.65) (derived similarly to (7.44) by using $\sigma(t) = 0$). By using the triangle inequality, the output estimation error $\epsilon_s^y(t)$ from (7.50) satisfies

$$\begin{aligned} |\epsilon_s^y(t)| &\geq |\epsilon_{s,2}^x(t)| - |\xi(t)| - |\hat{\sigma}(t)| \\ &\geq |\epsilon_{s,2}^x(t)| - \bar{\xi}_d - |\hat{\sigma}(t)|. \end{aligned} \tag{7.68}$$

For process fault identification, the inequality $|\epsilon_s^y(t)| > \bar{\epsilon}_s^y(t)$ (possibility of a sensor fault is excluded) must hold at some time $t > T_d$, so the sensor fault identification condition in (7.64) is obtained.

The second process fault identification condition given by (7.66) is obtained as follows. When a process fault occurs, and the fault is detected, the residual $r_s(t)$ from (7.61) becomes

$$r_s(t) = H_p(s)\big[A\epsilon_s^x(t) + \Delta f_s(t) + \Delta\eta_L(t) + \beta^x(t - T_0^x)\phi(x(t), u(t))\big] + \epsilon_\xi(t). \tag{7.69}$$

Using the triangle inequality, the k-th component of (7.69) satisfies

$$\begin{aligned} |r_s^{(k)}(t)| \geq\ & |H_p(s)\big[\beta^x(t - T_0^x)\phi^{(k)}(x(t), u(t))\big]| \\ & - |H_p(s)\big[A^{(k)}\epsilon_s^x(t) + \Delta f_s^{(k)}(t) + \Delta\eta_L^{(k)}(t)\big] + \epsilon_\xi^{(k)}(t)|. \end{aligned} \tag{7.70}$$

For process fault determination, the inequality $|r_s^{(k)}(t)| > \bar{r}_s^{(k)}(t)$ for some $k = 1, \ldots, n$ (possibility of a sensor fault is excluded) must hold at some time $t > T_d$, so the sensor fault identification condition in (7.66) is obtained.

(b) When a fault is detected at time T_d, one of the estimators initiated for identifying the fault type is the one for process faults given in Section 7.4.1, which satisfies $|r_p^{(k)}(t)| \leq \bar{r}_p^{(k)}(t)$ for all $k = 1, \ldots, n$ and $t > T_d + T_{L,2}$ if a process fault has occurred. So, when a sensor fault occurs, and the fault is detected, the residual $r_p(t)$ from (7.32), by using (7.14) and, by adding and subtracting $g(x(t) + \xi(t), u(t))$, becomes

$$\begin{aligned} r_p(t) = H_p(s)\big[& g(x(t), u(t)) - g(x(t) + \xi(t), u(t)) + g(x(t) + \xi(t), u(t)) \\ & - g(x(t) + \xi(t) + \sigma(t), u(t))\big] \\ & + H(s)\big[\sigma(t)\big] + H_p(s)\big[\Delta\eta_{L,2}(t)\big] - h_p(t)\xi(T_d) + \epsilon_\xi(t). \end{aligned} \tag{7.71}$$

By using the triangle inequality and Assumptions 7.6, 7.7, the k-th component of (7.71)

Christodoulos Keliris

satisfies

$$
\begin{aligned}
|r_p^{(k)}(t)| \geq & \left|H(s)\left[\sigma^{(k)}(t)\right] + H_p(s)\left[g^{(k)}(x(t)+\xi(t),u(t)) - g^{(k)}(x(t)+\xi(t)+\sigma(t),u(t))\right]\right| \\
& - |\epsilon_{\Delta g}^{(k)}(t)| - |H_p(s)\left[\Delta\eta_{L,2}^{(k)}(t)\right]| - |h_p(t)||\xi^{(k)}(T_d)| - |\epsilon_{\xi}^{(k)}(t)| \\
\geq & \left|H(s)\left[\sigma^{(k)}(t)\right] + H_p(s)\left[g^{(k)}(x(t)+\xi(t),u(t)) - g^{(k)}(x(t)+\xi(t)+\sigma(t),u(t))\right]\right| \\
& - \bar{\epsilon}_{\Delta g}^{(k)}(t) - \bar{H}_p(s)\left[\bar{\eta}_{L,2}^{(k)}(y(t),u(t))\right] - |h_p(t)|\bar{\xi}^{(k)} - \bar{\epsilon}_{\xi}^{(k)}(t) \\
\geq & \left|H(s)\left[\sigma^{(k)}(t)\right] + H_p(s)\left[g^{(k)}(x(t)+\xi(t),u(t)) - g^{(k)}(x(t)+\xi(t)+\sigma(t),u(t))\right]\right| \\
& - \bar{r}_p^{(k)}(t)
\end{aligned}
$$

and therefore, for sensor fault identification the inequality $|r_p^{(k)}(t)| > \bar{r}_p^{(k)}(t)$ (possibility of a process fault is excluded) must hold at some time $t > T_d + T_{L,2}$ for some $k = 1, 2, \ldots, n$, so the final sensor fault identification condition in (7.67) is obtained. □

7.6 Simulation Results

In this section, we apply the proposed approach to an example of single-link robotic arm with a revolute elastic joint with motion dynamics described by [74, 95]:

$$
\begin{aligned}
J_l\ddot{q}_l + F_l\dot{q}_l + k(q_l - q_m) + (m+\Delta m)gh\sin(q_l) &= 0 \\
J_m\ddot{q}_m + F_m\dot{q}_m - k(q_l - q_m) &= k_\tau u,
\end{aligned}
$$

where q_l is the angular position of the link, q_m is the angular position of the motor, J_l and J_m are the link and motor inertia respectively, F_l and F_m are the viscous coefficients of the link and motor respectively, k is the elastic constant, m is the link mass and Δm is the link mass inaccuracy, g is the gravity constant, h is the center of the mass, k_τ is the amplifier gain and u is the torque input delivered by the motor. The values of the parameters in SI units are: $J_l = 4.5$, $J_m = 1$, $F_l = 0.5$, $F_m = 1$, $k = 2$, $m = 4$, $\Delta m = 0.05m$, $g = 9.8$, $h = 0.5$, $k_\tau = 1$ and $u = 2\sin(0.25t)$. By selecting $x_1 = \dot{q}_m$, $x_2 = q_m$, $x_3 = \dot{q}_l$ and $x_4 = q_l$ the nonlinear uncertain model becomes

$$
\dot{x}(t) = Ax(t) + f(x_4(t), u(t)) + \eta(x_4(t)), \tag{7.72}
$$

Christodoulos Keliris

where $x(t) = [x_1(t), x_2(t), x_3(t), x_4(t)]^\top$ is the state vector,

$$A = \begin{bmatrix} -\frac{F_m}{Jm} & -\frac{k}{J_m} & 0 & \frac{k}{J_m} \\ 1 & 0 & 0 & 0 \\ 0 & \frac{k}{J_l} & -\frac{F_l}{J_l} & -\frac{k}{J_l} \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad f(x_4(t), u(t)) = \begin{bmatrix} \frac{k_\tau}{J_m} u(t) \\ 0 \\ -\frac{mgh}{J_l} \sin(x_4(t)) \\ 0 \end{bmatrix}$$

$$\eta(x_4(t)) = \begin{bmatrix} 0 \\ 0 \\ -\frac{\Delta mgh}{J_l} \sin(x_4(t)) \\ 0 \end{bmatrix}.$$

The initial conditions of the plant are set as $x(0) = 0$. The measurement noise $\xi^{(k)}(t)$ is generated from a uniform distribution with 5% uncertainty of the state. The nominal nonlinear function $f(x_4(t), u(t)$ satisfies the Lipschitz condition in $x_4(t) \in \mathbb{R}$ for all u, t with a Lipschitz constant $\lambda_f = mgh/J_l$. Note that, component-wise the function $f^{(k)}$ has a Lipschitz constant λ_{f_k}, $k = 1, \ldots, 4$ with $\lambda_{f_3} = mgh/J_l$ and $\lambda_{f_1} = \lambda_{f_2} = \lambda_{f_4} = 0$.

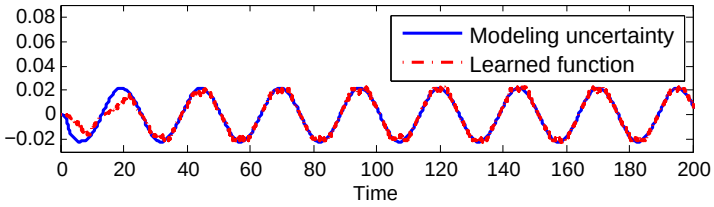
The filter used for dampening noise is $H(s) = sH_p(s)$ where $H_p(s) = \frac{50^2}{(s+10)(s+50)^2}$, and by using $\alpha = 10$, the required implementation filters are given by $H_1(s) = \frac{s}{s+10}$ and $H_2(s) = \frac{50^2}{(s+50)^2}$. Note that $H_p(s)$ has a non-negative impulse response and therefore the filter required for the threshold implementation is $\bar{H}_p(s) = H_p(s)$. For simplicity, the bounds on the filtered noise used are found from the simulation and are $\bar{\epsilon}_\xi^{(1)} = 0.01$, $\bar{\epsilon}_\xi^{(2)} = 0.04$, $\bar{\epsilon}_\xi^{(3)} = 0.002$, $\bar{\epsilon}_\xi^{(4)} = 0.006$. Similarly, the bounds on the filtered mismatch function used are $\bar{\epsilon}_{\Delta g}^{(1)} = 0.005$, $\bar{\epsilon}_{\Delta g}^{(2)} = 0.0005$, $\bar{\epsilon}_{\Delta g}^{(3)} = 0.002$ and $\bar{\epsilon}_{\Delta g}^{(4)} = 0.001$.

The objective is to apply the proposed fault diagnosis framework to achieve the following tasks: i) learn the modeling uncertainty η, ii) detect any faults, iii) identification of the type of the fault (process or sensor) and fault estimation. In addition, the enhancement of the fault detectability by using the integrated filtering and learning approach is also demonstrated.

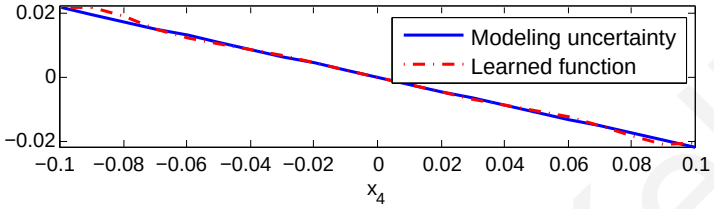
i) *Learning the modeling uncertainty*. This task is based on Sections 7.2.1 and 7.2.2. The training time to learn the modeling uncertainty is selected as $T_L = 200$ sec. We use a radial basis function network with 30 centers distributed evenly in the region of interest $x^{(4)} = [-0.25, 0.25]$ and constant variance $\sigma^2 = 0.02^2$ for each basis function. It is also considered that the radius of the hypersphere for the parameter adaptation is $M_p = 1$ and the constant matrix used for learning is $\Gamma^x = 2\mathbb{I}$, where $\mathbb{I}$ is the identity matrix. The results of the learning process are shown in Figure 7.3, in which the modeling uncertainty $\eta^{(3)}$ and its estimation $\hat{\eta}_L^{(3)}$ are shown. As it can be seen, the learning procedure is effective and the

Christodoulos Keliris

modeling uncertainty is learned quickly.



(a) Learning of the modeling uncertainty.



(b) Learned modeling uncertainty after the end of learning.

Figure 7.3: Learning of the modeling uncertainty.

ii) *Fault detection*. This task is based on Sections 7.2.3 and 7.2.4. After the end of the learning period at T_L, the fault detection scheme is initiated by implementing the residual and the detection threshold according to (7.19) and (7.24) respectively using a bound for the learned modeling uncertainty $\bar{\eta}_L = [0, 0, 0.05, 0]^\top$.

iii) *Fault type identification and estimation*. After the detection of the fault, the fault identification and estimation are initiated according to Section 7.4.1 to investigate the process fault occurrence and Section 7.4.2 to investigate the sensor fault occurrence. For the investigation of a process fault, we essentially resume the learning process for the modeling uncertainty in order to learn the combined effect of the modeling uncertainty and the potential process fault. The training time is selected as $T_{L,2} = 200$ sec and the learning is conducted during the time interval $[T_d, T_d + T_{L,2}]$. The matrix used for learning is changed to $\Gamma^x = 20\mathbb{I}$ to allow learning of potential faster changes and, the bound used for the learned modeling uncertainty combined with the process fault is the same as before, i.e. $\bar{\eta}_{L,2} = \bar{\eta}_L$. After, the end of the training, if a residual $r_p^{(k)}$ (7.31) exceeds its corresponding process fault isolation threshold $\hat{r}_p^{(k)}$ (7.33) for some $k = 1, \ldots, 4$, then the occurrence of a process fault is excluded and it is guaranteed that a sensor fault has occurred. Once a process fault has been excluded, then the sensor fault estimation is given by the $\hat{\sigma}(t)$ which contains the estimation of each sensor fault.

For the investigation of a sensor fault, we implement the estimation model (7.34)-(7.36). The gain matrix Λ is selected via pole placement so that the eigenvalues of A_0 are located

Christodoulos Keliris

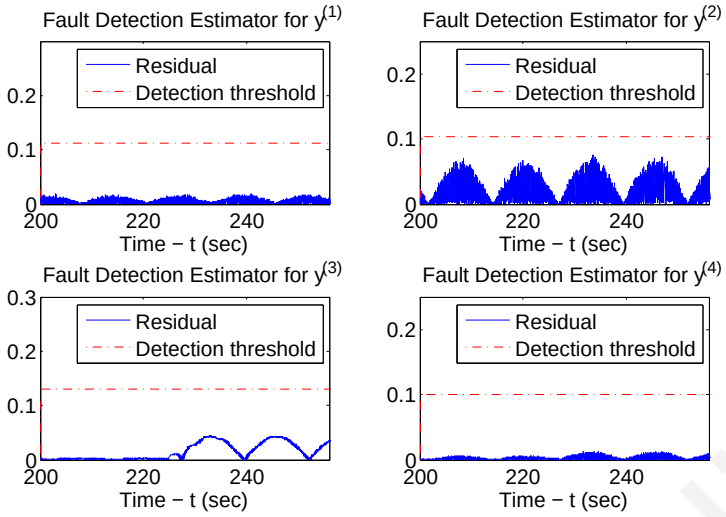


Figure 7.4: Process fault detection without the use of learning or filtering.

at $[-9.1, -9.8, -10.5, -11.2]$. The residual in this case is implemented according to (7.58), (7.59) and the sensor isolation thresholds according to (7.63), (7.49), by using $\mu = 1$, $\rho = 7.5$, $M_\sigma = 0.1$ and $\Gamma^y = 0.1$. If at some time $t > T_d$ a residual $r_s^{(k)}$ (7.58) exceeds its corresponding sensor fault isolation threshold $\hat{r}_s^{(k)}$ (7.63) for some $k = 1, \ldots, 4$, then the possibility of a sensor fault is excluded and it is guaranteed that a process fault has occurred. The estimation of the process fault including the modeling uncertainty is given by $\hat{\eta}_{L,2}$ for $t > T_d + T_{L_2}$ (after the end of learning).

In the sequel we will examine two cases; the first one considers the occurrence of a process fault and the second one of a sensor fault.

Case 1 - Process fault. The process fault is considered to be a fault that results in the reduction of the mass of the link, which occurs abruptly at time $T_0^x = 225$ sec. Specifically the fault function is given by $\phi(x(t), u(t) = [0, 0, -\theta_x \frac{(m+\Delta m)gh}{J_l} \sin(x^{(4)}(t)), 0]^\top$, where $\theta_x \in [-1, 0]$ represents the percentage change of the mass, which in this example is considered $\theta_x = -0.5$. At first, in order to demonstrate effectiveness of the proposed scheme, we obtain the fault detection results without the use of filtering or learning. These results, which are shown in Figure 7.4, are obtained by using suitable bounds on the noise magnitude and modeling uncertainty in order to guarantee no false alarms as described in Remark 7.2. As it can be seen from Figure 7.4, the detection thresholds are too conservative to be crossed by their respective residuals and hence, the fault is not detected. On the other hand, by using the integrated filtering and learning approach proposed in this work, we have successful fault detection, as it can be seen from Figure 7.5 which shows the residual $r^{(k)}$ and corresponding detection threshold $\bar{r}^{(k)}$ for each measurement $y^{(k)}$, $k = 1, \ldots, 4$. Specifically, the residual

Christodoulos Keliris

exceeds its threshold in the case of the estimator that monitors $y^{(3)}$ at about $T_d = 225.5$ sec, and hence the fault is detected.

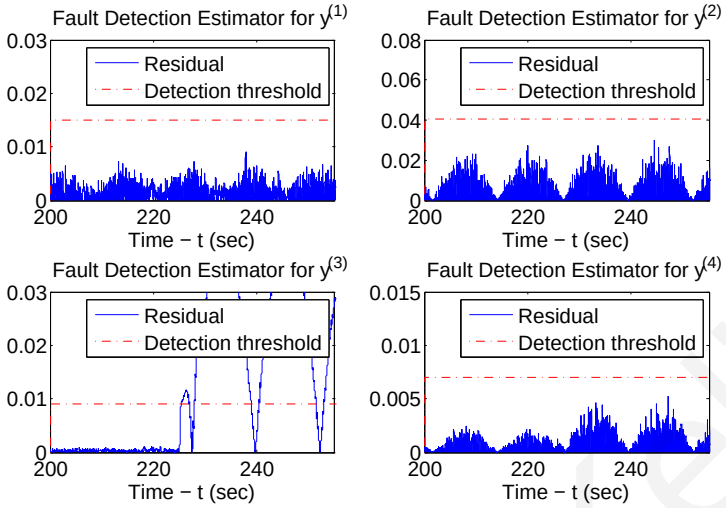
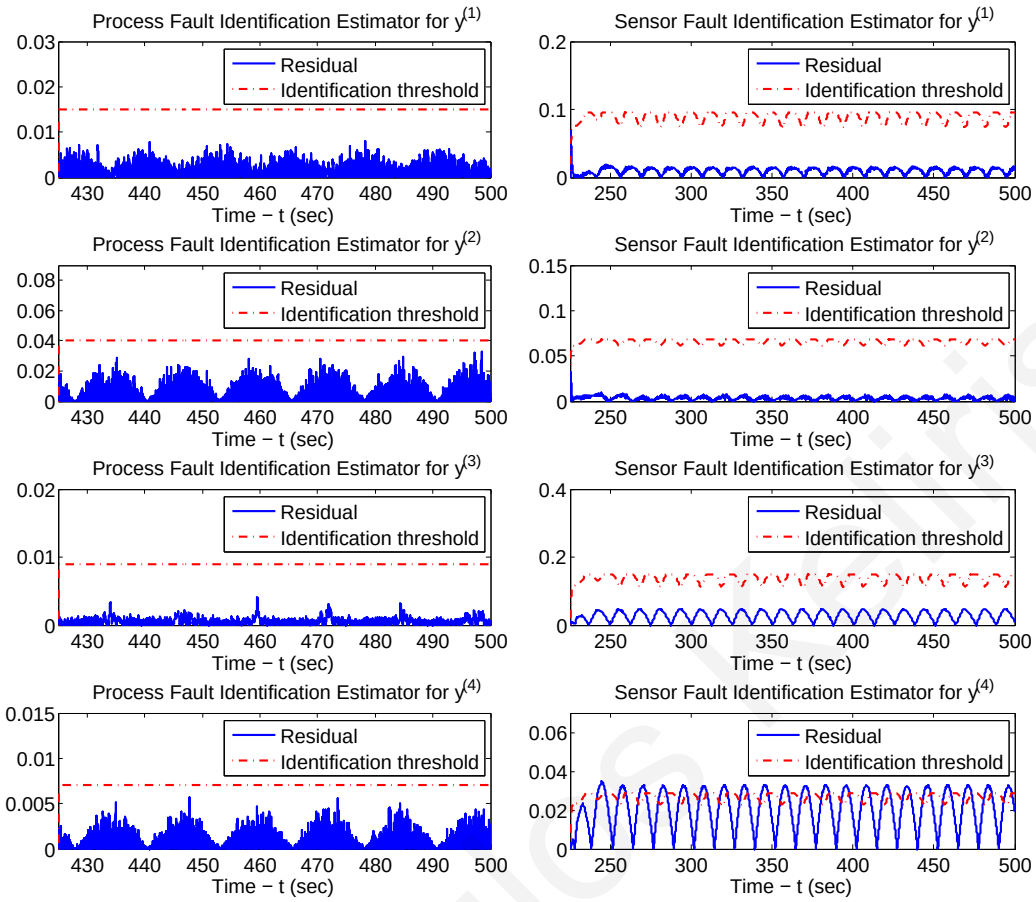


Figure 7.5: Process fault detection with the integrated filtering and learning approach.

After the fault is detected, the process and sensor fault estimation models are initiated and, their results are illustrated in Figures 7.6a and 7.6b respectively. As it can be seen from Figure 7.6b, in FIE for $y^{(4)}$ the residual exceeds its sensor identification threshold at about 241 sec and hence the case of a sensor fault is excluded and, the occurrence of a process fault is concluded. In addition, note that in Figure 7.6a all the residuals remain below their corresponding process fault identification thresholds after the end of the learning. The process fault including the modeling uncertainty $\eta^{(3)} + \phi^{(3)}$ and its estimation $\hat{\eta}^{(3)}_{L,2}$ are shown in Figure 7.7, where it is seen that the function has been learned very well.

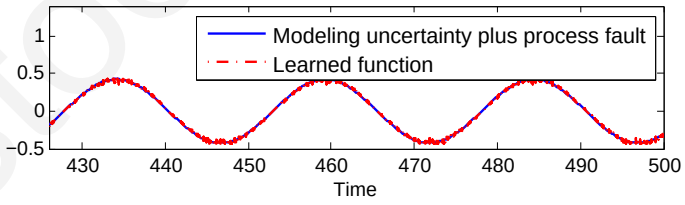
Christodoulos Keliris



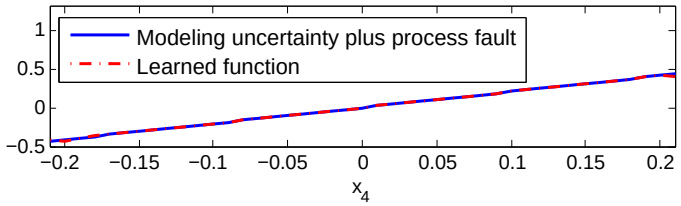
(a) FIEs for Process fault.

(b) FIEs for Sensor fault.

Figure 7.6: Fault Identification Estimators for Process and Sensor faults in the occurrence of a **process** fault.



(a) Learning of process fault including modeling uncertainty.



(b) Learned process fault including modeling uncertainty after the end of learning.

Figure 7.7: Learning of process fault including modeling uncertainty.

Christodoulos Keliris

Case 2 - Sensor fault. At time $T_0^y = 225$ sec two sensors (2 and 3) become faulty by measuring the state variables $x^{(2)}$ and $x^{(3)}$ with a constant sensor bias $\sigma^{(2)} = 0.07$ and $\sigma^{(3)} = -0.07$ respectively (in addition to noise). Figure 7.8 shows the residual $r^{(k)}$ and corresponding detection threshold $\bar{r}^{(k)}$ for each measurement $y^{(k)}$, $k = 1, \ldots, 4$. As it can be seen from Figure 7.8, the residual exceeds its threshold in the case of the estimators that monitor $y^{(1)}$ and $y^{(4)}$. The fault is detected the first time the threshold is exceeded at time $T_d = 226.5$ sec.

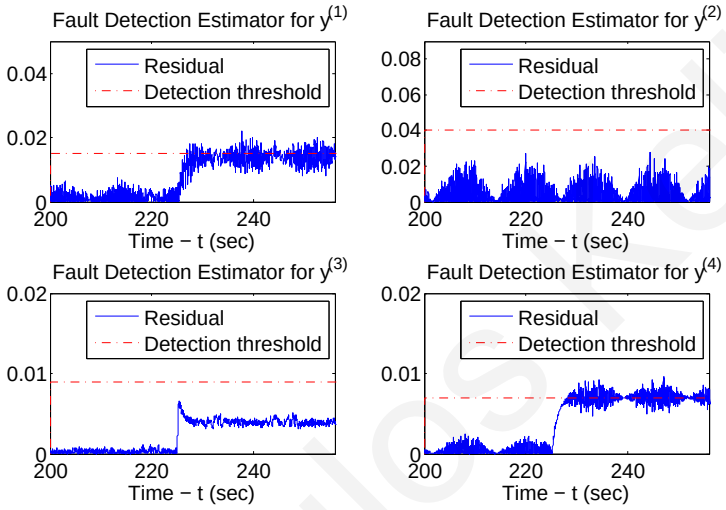
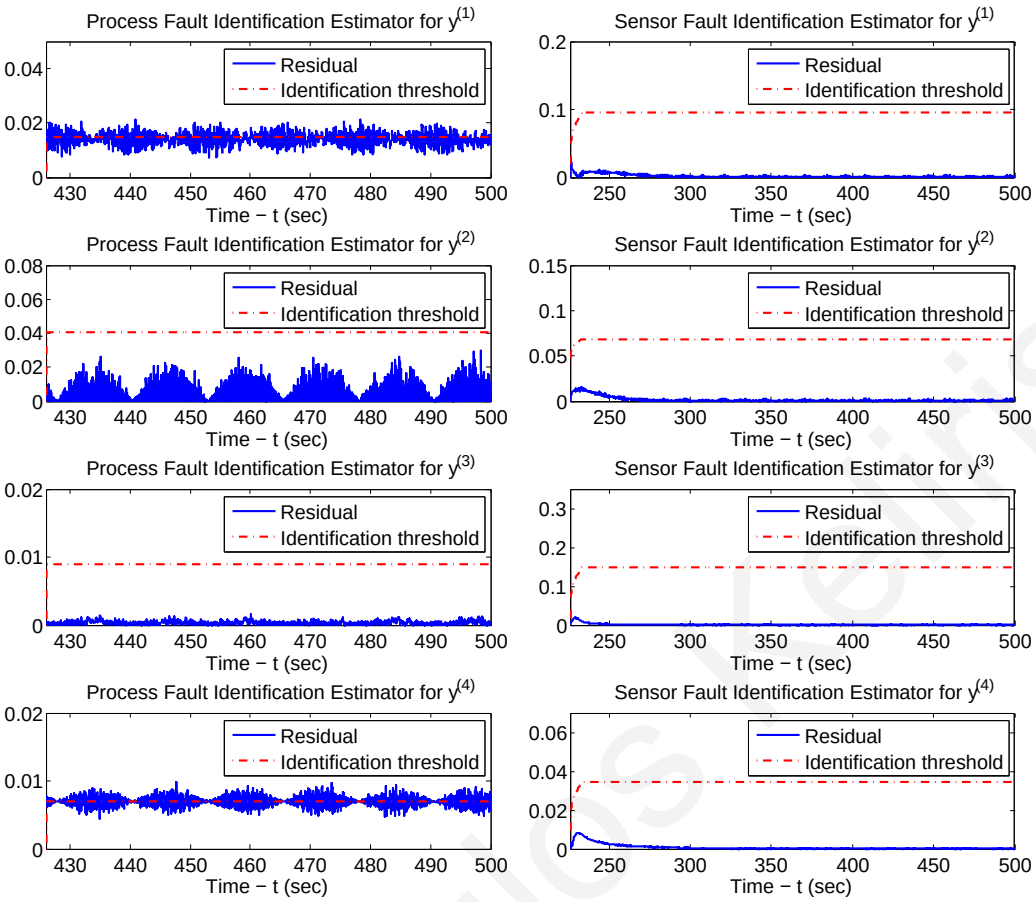


Figure 7.8: Sensor fault detection.

After the fault is detected, the process and sensor fault estimation models are initiated and, their results are illustrated in Figures 7.9a and 7.9b respectively. As it can be seen from Figure 7.9b, all the residuals remain below their sensor identification thresholds, whereas in Figure 7.9a the residuals of the FIEs that monitor $y^{(1)}$ and $y^{(4)}$ exceed their corresponding process fault identification thresholds immediately after the end of the training phase at time $T_d + T_{L,2} = 425$ sec. Therefore, the case of a process fault is excluded and, the occurrence of a sensor fault is concluded. The estimation of the sensor fault $\hat{\sigma}(t)$ is shown in Figure 7.10, in which it can be seen that the sensor fault is estimated correctly.

Christodoulos Keliris



(a) FIEs for Process fault. (b) FIEs for Sensor fault.

Figure 7.9: Fault Identification Estimators for Process and Sensor faults in the occurrence of a **sensor** fault.

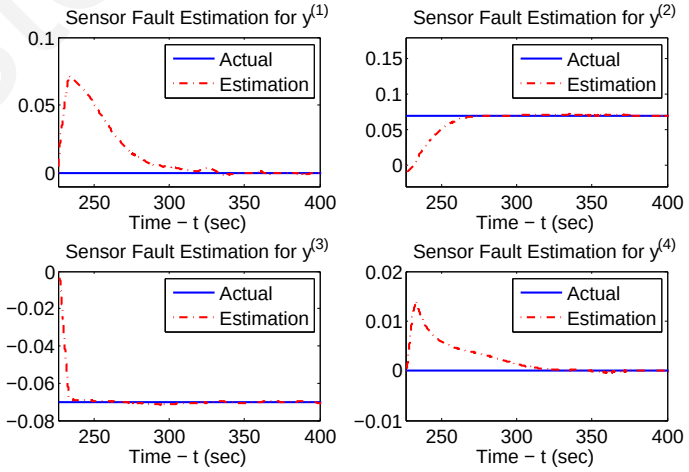


Figure 7.10: Learned sensor fault.

Christodoulos Peiris

7.7 Conclusion

In this chapter a unified fault diagnosis approach for process and sensor faults in a class of continuous-time, nonlinear systems with modeling uncertainties and measurement noise is presented, which integrates learning with filtering techniques for obtaining tight detection thresholds. The scheme exploits the filtering framework to deal with the learning and noise attenuation tasks in a two stage filtering process for designing the required signals both tasks. By using the learned function in the design of the residual signal and exploiting the noise attenuation of a general class of filters, enhanced fault detectability is achieved. When a fault is detected two estimation models, one for process and one for sensor faults, are initiated in order to learn the potential fault that has occurred. Through, the design of suitable residual and identification thresholds, the fault type can be determined and the estimation of the fault is obtained.

Christodoulos Kellis

Chapter 8

Conclusions and Future Work

8.1 Contributions

This dissertation investigates the problem of fault detection in nonlinear uncertain systems with noisy measurements and, proposes a fault detection scheme using filtering under a deterministic framework to counteract the noise effects, so that enhanced fault detection is achieved. A comprehensive study of the problem under a wide spectrum of cases including continuous/discrete time and by considering full state measurement/input-output cases was undertaken and, a novel fault detection filtering approach was devised. The key novelty and main contribution of the proposed work is the integration of a general class of filters into the design of the residual and threshold signals in a way that takes advantage of the filtering noise suppression properties and as a result, a robust fault detection scheme is designed. Essentially, filtering dampens the effect of measurement noise in a certain frequency range allowing to set tighter detection thresholds and thus enhancing fault detectability. In all cases, the adaptive fault detection thresholds are obtained under a rigorous analytical framework guaranteeing no false alarms and, respective fault detectability conditions characterizing quantitatively a class of detectable faults are derived. The main implications of the filtering novelty is rigorously investigated in the continuous-time framework with regards to the impact of the filters' poles and order on the fault detection time (Chapter 3).

The case in which the systems under consideration belong to a class of input-output systems is investigated in depth (Chapter 4). In the general case, nonlinear observer design is used to generate the residual signals required for fault detection and then, a distributed fault detection scheme and the corresponding adaptive thresholds are designed based on the observer characteristics. At the same time, filtering is used in order to attenuate the effect of

Christodoulos Keliris

measurement noise, which facilitates less conservative thresholds and enhanced robustness. A particular nonlinear observer design method is proposed, although it must be stressed that any nonlinear observer design satisfying a given condition can also be used. A simplified problem formulation is also investigated as a special case, in which the nonlinear observer design is not needed and therefore a simplified estimation model is formulated and a modified distributed fault detection filtering scheme is proposed. Finally, the discrete-time variants of the proposed fault filtering scheme are also designed (Chapters 5 and 6).

An additional important contribution of this work, is the investigation of the particular design of the distributed fault detection scheme that makes use of the measurements of the interconnection variables which is considered in Chapters 3-6. As it is shown in Chapter 6, the design possesses important characteristics regarding fault propagation among the interconnected subsystems, allowing the design of a high-level isolation scheme. The distributed fault detection scheme is based on local fault filtering agents with each one assigned to monitor one subsystem. Each local fault detection agent receives the input and output measurements of the subsystem it monitors and also the output measurements of all the interconnected subsystems that influence the subsystem under consideration. Finally, the local fault detection scheme provides a decision regarding the health of the subsystem it monitors. The main characteristic of the distributed fault diagnosis scheme is that a process fault occurring in one subsystem can only be detected by its corresponding detection agent whereas a sensor fault in a subsystem can be detected by either its corresponding detection agent or the detection agent of another subsystem that is affected by the subsystem where the sensor fault occurred. This discriminating factor allows the extrapolation of useful information regarding the fault that has occurred.

The final major contribution of this work, is the design of a unified framework for fault diagnosis for a class of nonlinear uncertain systems with noisy measurements that deals with both process and sensor faults (Chapter 7). The proposed scheme integrates learning with filtering techniques in a unified framework and allows the derivation of tight detection thresholds. This is accomplished in two ways: at first by learning the modeling uncertainty through adaptive approximation methods, so that the learned function is used for the derivation of the residual signal, and then by using filtering for dampening measurement noise. The required signals for both tasks are derived by exploiting the properties of the filtering framework through a two-stage filtering process and allows the decoupling of the learning and noise dampening tasks, into two separate and independent tasks. As a result, even tighter thresholds can be achieved, thus enhancing fault detectability. Finally, by using learning methods

Christodoulos Kallis

the type of the fault can be identified (process or sensor) and its estimation is generated so that it can be used for fault accommodation purposes.

8.2 Concluding Remarks and Future Work

Below we discuss some issues concerning some aspects of the proposed scheme that have not been elaborated so far and can provide useful and insightful information regarding the approach and its applicability in real applications.

8.2.1 Effect of the feedback controller

Feedback controllers are used for controlling the plant and, the control law implemented by the feedback controllers makes use of the sensor measurements. In the proposed scheme, the local (feedback) controllers u_I do not create any negative consequences on the fault detection aspect. For instance, let us consider the case of an actuator fault. Since the actuator used in the estimation model is implemented in the software of the fault detection agent, hence, any discrepancy between the actual actuator output acting on the system and the nominal actuator output implemented in the detection agent can only aid and enhance the detection of the actuator fault. In the case of a sensor fault, the faulty measurements are used by the real actuation law acting on the system and by the nominal actuator law acting on the estimation model. As a result, no discrepancies are introduced in the actuation law in order to cause a potential detection of the sensor fault. In this case, the fault detection is based on the discrepancies, created by the sensor faults, between the state variables of the system and the estimation model. Finally, in the case of a process fault which affects the system states, the measurements of these variables contain the process fault effects and, similarly to the sensor fault case, the use of feedback controllers does not affect the fault detection scheme.

In this work, due to the assumption that during the learning period no faults occur, the feedback controller does not affect the performance of the learning methodology for the modeling uncertainty. This though, is not true in general. If faults are allowed during the learning phase, then potential sensor/actuator faults could deteriorate significantly the learning capabilities of the adaptive approximation procedure, since wrong information will be used to excite directly and indirectly (through the feedback controller) the basis functions. In the event of a process fault though, the performance of the learning methodology is not affected, since in this case the learning procedure adapts in order to learn the combined effect

Christodoulos Keliris

of the modeling uncertainty with the process fault, something that is exploited in the design of the unified fault diagnosis approach described in Section 7.4.1.

8.2.2 Scalability and distributed aspects of the proposed method

The proposed distributed fault diagnosis approach relies on the availability of the measurements of the interconnection variables among subsystems, so that these can be communicated among the fault detection agents. This way, the structure of the distributed fault diagnosis mirrors exactly the structure of the interconnected nonlinear systems (i.e. see Figure 3.1 or 4.1). Therefore, the distributed nature of the proposed fault diagnosis approach stems from the communication of (some) of the measurements between the fault detection agents. As a result, the proposed approach has inherent expandability characteristics, by allowing the addition or removal of subsystems with their respective fault detection agents, without causing any issues in the remaining system and its fault detection mechanism. This is particularly useful, since any subsystem can be accompanied by its respective fault detection agent and can be used in a "plug and play" approach in other implementations without the need for reworking the fault detection module.

8.2.3 Applicability of the approach in real systems and large-scale applications

The simulation examples used in this work were mainly of academic nature, in order to illustrate the effectiveness of the proposed approach and emphasize the specific contribution of the scheme in addressing the problems in each chapter. As a result, the applicability of the approach in a real system has not been investigated and, it would be interesting to test the scheme in real physical systems, such as the three-tank system benchmark, something that would allow the performance comparison of the proposed approach and other proposed schemes available in the literature. The filtering novelty can aid in the detection of faults in a specific frequency range of the system operation that is of interest, or banks of filters can be used to monitor various ranges in the frequency domain.

The proposed approach can also be exploited in large-scale applications, in which the large-scale system is distributed and, the multiple subsystems that constitute it span across different physical locations. In such cases, relying in a centralized fault diagnosis approach in which all the information is gathered and processed centrally not only poses a significant single-point of failure risk but also, requires more computation power and communi-

Christodoulos Keliris

cation bandwidth. Therefore, the proposed approach can provide a natural fault diagnosis framework that is easy to maintain, add and remove subsystems with their corresponding detection modules. Of course, real world systems create additional issues that need to be addressed first for making the proposed approach applicable: i) the physical system works under continuous-time, whereas the implementation of the fault diagnosis approach requires the execution in discrete-time and hence a suitable sampling and synchronization strategy must be devised and, ii) the communicated interconnection variables are considered synchronized and available whereas in real systems, communication delays or packet losses can occur.

8.2.4 Future Work

Potential future work can focus on the following topics/issues:

- **Comprehensive fault isolation methodology.** The proposed fault diagnosis scheme encompasses important fault isolation characteristics that can be further exploited. Therefore, a possible future work would be the design of a comprehensive fault diagnosis scheme, utilizing filtering and adaptive approximation techniques, so that the particular fault that has occurred can be determined among a class of possible faults. In this case, the dynamics of each potential fault can be considered as the basis functions of the adaptive approximation procedure, so that the identification of the specific fault that has occurred and the determination of its magnitude can be achieved (fault isolation and identification).

- **Filter selection.** In this work, the primary task for filter use is the attenuation of noise and its effects, so that tighter thresholds can be obtained. The detectability analysis conducted, provides some guidelines for filter selection. Therefore, a possible future work would be the derivation of filter selection criteria for achieving particular performance goals. The proposed method is very promising in enhancing the performance of other proposed methods in the literature, since the treatment of the filtering design as a linear state transformation maintains the structure of the filtered system dynamics with that of the actual system. Therefore, the extra degree of freedom provided by the use of filtering can improve the overall performance and robustness of other schemes.

- **Extension to unknown interconnection variables.** The proposed distributed fault diagnosis scheme proposed in this work relies on the availability of the interconnec-

Christodoulos Kelinis

tion variables, so that the measurements of these interconnection variables can be exchanged among the fault detection agents. Therefore, a natural extension of this work would be the design of a distributed fault diagnosis scheme in the case when the interconnection variables are not measurable, which poses a more challenging task. A glimpse of the difficulties incurred in this task is given in Chapter 4, in which unmeasurable state variables in the local nominal system nonlinearity were considered and successfully tackled through nonlinear observer design. In the more general case of unmeasurable interconnection variables, advanced control theory will be required in order to design suitable local observers that exchange the estimates of the interconnection variables and guarantee the stability of the overall estimation scheme.

Christodoulos Kellis

Bibliography

[1] K. Adjallah, D. Maquin, and J. Ragot, "Nonlinear observer-based fault detection," in *IEEE Conference on Control Applications*, no. 3, 1994, pp. 1115–1120.

[2] M. Basseville and I. Nikiforov, *Detection of abrupt changes: theory and application*. Prentice-Hall, 1993.

[3] G. Besançon, Ed., *Nonlinear Observers and Applications*. Springer, 2007.

[4] M. Blanke, M. Kinnaert, J. Lunze, and M. Staroswiecki, *Diagnosis and Fault-Tolerant Control*, 2nd ed. Springer Verlag, 2010.

[5] F. Boem, R. M. G. Ferrari, C. Keliris, T. Parisini, and M. M. Polycarpou, "A Distributed Networked Approach for Fault Detection of Large-scale Systems,," *IEEE Transactions on Automatic Control*, (under review).

[6] F. Boem, R. M. Ferrari, and T. Parisini, "Distributed Fault Detection and Isolation of Continuous-Time Nonlinear Systems," *European Journal of Control*, vol. 5-6, pp. 603–620, 2011.

[7] W. Brogan, *Modern Control Theory*. Prentice-Hall, Englewood Cliffs, NJ, 1985.

[8] F. Callier and C. Desoer, *Linear system theory*, 1st ed. Springer, 1994.

[9] A. Castillo and P. J. Zufiria, "Fault Detection Schemes for Continuous-Time Stochastic Dynamical Systems," *IEEE Transactions on Automatic Control*, vol. 54, no. 8, pp. 1820–1836, Aug. 2009.

[10] J. Chen and R. J. Patton, *Robust Model-Based Fault Diagnosis for Dynamic Systems*. Kluwer Academic Publishers Norwell, MA, USA, 1999.

[11] R. H. Chen and J. L. Speyer, "A generalized least-squares fault detection filter," *International Journal of Adaptive Control and Signal Processing*, vol. 14, no. 7, pp. 747–757, Nov. 2000.

[12] ——, "Robust multiple-fault detection filter," *International Journal of Robust and Nonlinear Control*, vol. 12, no. 8, pp. 675–696, Jul. 2002.

[13] T. Chen, C. Wang, and D. Hill, "Rapid Oscillation Fault Detection and Isolation for Distributed Systems via Deterministic Learning," *Neural Networks and Learning Systems, IEEE Transactions on*, vol. 25, no. 6, pp. 1187–1199, 2014.

[14] W. Chen and M. Saif, "Observer-based strategies for actuator fault detection, isolation and estimation for certain class of uncertain nonlinear systems," *Control Theory & Applications, IET*, vol. 1, no. 6, pp. 1672–1680, 2007.

Christodoulos Keliris

[15] H. Cheng and S. Yau, "More explicit formulas for the matrix exponential," *Linear algebra and its applications*, vol. 163, no. 1997, pp. 131–163, 1997.

[16] W. Chung and J. L. Speyer, "A game theoretic fault detection filter," *IEEE Transactions on Automatic Control*, vol. 43, no. 2, pp. 143–161, 1998.

[17] R. Clark, "Instrument fault detection," *IEEE Transactions on Aerospace and Electronic Systems*, no. 3, pp. 456–465, 1978.

[18] R. Da and C. Lin, "Sensor failure detection with a bank of Kalman filters," in *American Control Conference, 1995. Proceedings of the*, vol. 2, 1995, pp. 1122–1126.

[19] C. De Persis and A. Isidori, "A geometric approach to nonlinear fault detection and isolation," *IEEE Transactions on Automatic Control*, vol. 46, no. 6, pp. 853–865, Jun. 2001.

[20] ——, "On the design of fault detection filters with game-theoretic-optimal sensitivity," *International Journal of Robust and Nonlinear Control*, vol. 12, no. 8, pp. 729–747, 2002.

[21] C. Desoer and M. Vidyasagar, *Feedback Systems: Input-Output Properties*, 1st ed. Academic Press, 1975.

[22] R. Dunia and S. Joe Qin, "Joint diagnosis of process and sensor faults using principal component analysis," *Control Engineering Practice*, vol. 6, no. 4, pp. 457–469, 1998.

[23] A. Emami-Naeini, M. Akhter, and S. Rock, "Effect of model uncertainty on failure detection: the threshold selector," *IEEE Transactions on Automatic Control*, vol. 33, no. 12, pp. 1106–1115, 1988.

[24] J. Farrell and M. M. Polycarpou, *Adaptive approximation based control: unifying neural, fuzzy and traditional adaptive approximation approaches*. Wiley-Blackwell, 2006.

[25] H. Ferdowsi, S. Jagannathan, and M. Zawodniok, "An online outlier identification and removal scheme for improving fault detection performance." *IEEE Transactions on Neural Networks*, vol. 25, no. 5, pp. 908–19, May 2014.

[26] H. Ferdowsi, D. Raja, and S. Jagannathan, "A decentralized fault prognosis scheme for nonlinear interconnected discrete-time systems," in *American Control Conference*, 2012, pp. 5900–5905.

[27] R. M. Ferrari, T. Parisini, and M. M. Polycarpou, "A robust fault detection and isolation scheme for a class of uncertain input-output discrete-time nonlinear systems," in *American Control Conference*, 2008, pp. 2804–2809.

[28] ——, "Distributed fault diagnosis with overlapping decompositions: an adaptive approximation approach," *IEEE Transactions on Automatic Control*, vol. 54, no. 4, pp. 794–799, 2009.

[29] ——, "Distributed fault detection and isolation of large-scale nonlinear systems: an adaptive approximation approach," *IEEE Transactions on Automatic Control*, vol. 57, no. 2, pp. 275–290, 2012.

Christodoulos Keliris

[30] T. Floquet, J. Barbot, W. Perruquetti, and M. Djemai, "On the robust fault detection via a sliding mode disturbance observer," *International Journal of Control*, vol. 77, no. 7, pp. 622–629, May 2004.

[31] P. Frank, "Fault diagnosis in dynamic systems via state estimation-a survey," *System fault diagnostics, reliability and related knowledge-based approaches*, vol. 1, pp. 35–98, 1987.

[32] ——, "Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy: A survey and some new results," *Automatica*, vol. 26, no. 3, pp. 459–474, 1990.

[33] P. Frank and S. X. Ding, "Survey of robust residual generation and evaluation methods in observer-based fault detection systems," *Journal of Process Control*, no. 6, pp. 403–424, 1997.

[34] P. Frank, G. Schrier, and E. Garcia, "Nonlinear observers for fault detection and isolation," in *New Directions in nonlinear observer design*, ser. Lecture Notes in Control and Information Sciences, H. Nijmeijer and T. Fossen, Eds. London: Springer, 1999, vol. 244, pp. 399–422.

[35] E. A. Garcia and P. Frank, "Deterministic nonlinear observer-based approaches to fault diagnosis: a survey," *Control Engineering Practice*, vol. 5, no. 5, pp. 663–670, 1997.

[36] J. Gauthier and I. Kupka, "Observability and observers for nonlinear systems," *SIAM Journal on Control and Optimization*, vol. 32, no. 4, pp. 975–994, 1994.

[37] J. Gertler, "Survey of model-based failure detection and isolation in complex plants," *IEEE Control Systems Magazine*, vol. 8, no. 6, pp. 3–11, 1988.

[38] ——, "Fault detection and isolation using parity relations," *Control Engineering Practice*, vol. 5, no. 5, pp. 653–661, May 1997.

[39] ——, *Fault detection and diagnosis in engineering systems*, 1st ed. CRC Press, 1998.

[40] H. Hammouri, P. Kabore, and M. Kinnaert, "A geometric approach to fault detection and isolation for bilinear systems," *IEEE Transactions on Automatic Control*, vol. 46, no. 9, pp. 1451–1455, 2001.

[41] H. Hammouri, M. Kinnaert, and E. El Yaagoubi, "Observer-based approach to fault detection and isolation for nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 44, no. 10, pp. 1879–1884, 1999.

[42] I. Hwang, S. Kim, Y. Kim, and C. E. Seah, "A Survey of Fault Detection, Isolation, and Reconfiguration Methods," *IEEE Transactions on Control Systems Technology*, vol. 18, no. 3, pp. 636–653, May 2010.

[43] P. Ioannou and J. Sun, *Robust adaptive control*. Prentice Hall, 1996.

[44] R. Isermann, "Process fault detection based on modeling and estimation methods - A survey," *Automatica*, vol. 20, no. 4, pp. 387–404, Jul. 1984.

[45] ——, "Supervision, fault-detection and fault-diagnosis methods - An introduction," *Control engineering practice*, vol. 5, no. 5, pp. 639–652, 1997.

Christodoulos Keliris

[46] ——, *Fault Diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance*, 1st ed. Springer, 2005.

[47] N. Kazantzis and C. Kravaris, "Nonlinear observer design using Lyapunov's auxiliary theorem," *Systems & Control Letters*, vol. 34, no. 5, pp. 241–247, 1998.

[48] C. Keliris and M. M. Polycarpou, "A Distributed Fault Detection Filtering Approach for a Class of Interconnected Continuous-Time Nonlinear Systems," in *Proc. of 50th IEEE Conference on Decision and Control and European Control Conference*, 2011, pp. 89–94.

[49] C. Keliris, M. M. Polycarpou, and T. Parisini, "A Distributed Fault Detection Filtering Approach for a Class of Interconnected Continuous-Time Nonlinear Systems," *IEEE Transactions on Automatic Control*, vol. 58, no. 8, pp. 2032–2047, 2013.

[50] ——, "A Distributed Fault Detection Filtering Approach for a Class of Interconnected Input-Output Nonlinear Systems," in *Proc. of European Control Conference*, 2013, pp. 422–427.

[51] ——, "A Distributed Fault Diagnosis Approach Utilizing Adaptive Approximation for a Class of Interconnected Continuous-Time Nonlinear Systems," in *Proc. of Control and Decision Conference*, 2014, pp. 6536–6541.

[52] ——, "A Robust Nonlinear Observer-based Approach for Distributed Fault Detection of Input-Output Interconnected Systems," *Automatica*, vol. 53, no. 3, pp. 408–415, 2015.

[53] ——, "Distributed Fault Diagnosis for Process and Sensor Faults in a Class of Interconnected Input-Output Nonlinear Discrete-Time Systems," *International Journal of Control*, 2015.

[54] ——, "A Unified Fault Diagnosis Approach Utilizing Filtering and Adaptive Approximation for Process and Sensor Faults in a Class of Continuous-Time Nonlinear Systems," *IEEE Transactions on Neural Networks and Learning Systems*, (to be submitted).

[55] S. Klinkhieo and R. J. Patton, "A Two-Level Approach to Fault-Tolerant Control of Distributed Systems Based on the Sliding Mode," in *7th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes, Barcelona, Spain*, 2009, pp. 1043–1048.

[56] A. Krener and A. Isidori, "Linearization by output injection and nonlinear observers," *Systems & Control Letters*, vol. 3, no. 1, pp. 47–52, 1983.

[57] N. Léchevin and C. Rabbath, "Decentralized Detection of a Class of Non-Abrupt Faults With Application to Formations of Unmanned Airships," *IEEE Transactions on Control Systems Technology*, vol. 17, no. 2, pp. 484–493, 2009.

[58] S. Leonhardt, "Methods of fault diagnosis," *Control Engineering Practice*, vol. 5, no. 5, pp. 683–692, May 1997.

[59] Y. Liu and P. Bauer, "Sufficient conditions for non-negative impulse response of arbitrary-order systems," in *IEEE Asia Pacific Conference on Circuits and Systems*, 2008, pp. 1410–1413.

Christodoulos Keliris

[60] K. Loparo, Z. Roth, and S. Eckert, "Nonlinear filtering for systems with random structure," *Automatic Control, IEEE*, vol. 31, no. 11, pp. 1064–1068, Nov. 1986.

[61] M. Massoumnia, G. Verghese, and A. Willsky, "Failure detection and identification," *IEEE Transactions on Automatic Control*, vol. 34, no. 3, pp. 316–321, 1989.

[62] P. Maybeck and P. Hanlon, "Performance enhancement of a multiple model adaptive estimator," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 31, no. 4, pp. 1240–1254, 1995.

[63] F. Olver, D. Lozier, R. Boisvert, and C. Clark, *NIST Handbook of Mathematical Functions*. Cambridge University Press, 2010.

[64] R. Patton and P. M. Frank, *Fault diagnosis in dynamic systems: theory and applications*. Prentice Hall, 1989.

[65] R. J. Patton, P. Frank, and R. Clarke, *Fault diagnosis in dynamic systems: theory and application*. Prentice Hall, 1989.

[66] R. J. Patton, C. Kambhampati, A. Casavola, P. Zhang, S. X. Ding, and D. Sauter, "A generic strategy for fault-tolerance in control systems distributed over a network," *European Journal of Control*, vol. 13, no. 2-3, pp. 280–296, 2007.

[67] A. Pertew, H. Marquez, and Q. Zhao, "Design of unknown input observers for Lipschitz nonlinear systems," *American Control Conference*, pp. 4198–4203, 2005.

[68] G. Phanomchoeng and R. Rajamani, "Observer design for Lipschitz nonlinear systems using Riccati equations," in *American Control Conference*, 2010, pp. 6060–6065.

[69] M. M. Polycarpou and A. Helmicki, "Automated fault detection and accommodation: A learning systems approach," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 25, no. 11, pp. 1447 – 1458, 1995.

[70] M. M. Polycarpou and A. Trunov, "Learning approach to nonlinear fault diagnosis: detectability analysis," *IEEE Transactions on Automatic Control*, vol. 45, no. 4, pp. 806–812, Apr. 2000.

[71] M. Powell, "The theory of radial basis function approximation in 1990," in *Advances in Numerical Analysis, Vol. II: Wavelets, Subdivision Algorithms, and Radial Basis Functions*, W. Light, Ed. Oxford, UK: Oxford University Press, 1992, vol. II, pp. 105–210.

[72] R. Rajamani, "Observers for Lipschitz nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 43, no. 3, pp. 397–401, Nov. 1998.

[73] R. Rajamani and A. Ganguli, "Sensor fault diagnostics for a class of non-linear systems using linear matrix inequalities," *International Journal of Control*, vol. 77, no. 10, pp. 920–930, 2004.

[74] V. Reppa, M. Polycarpou, and C. Panayiotou, "Adaptive Approximation for Multiple Sensor Fault Detection and Isolation of Nonlinear Uncertain Systems," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 1, pp. 137–153, 2014.

[75] K. Salahshoor, M. Mosallaei, and M. Bayat, "Centralized and decentralized process and sensor fault monitoring using data fusion based on adaptive extended Kalman filter algorithm," *Measurement*, vol. 41, no. 10, pp. 1059–1076, 2008.

Christodoulos Keliris

[76] B. Song and J. Hedrick, "Nonlinear Observer Design for Lipschitz Nonlinear Systems," in *Proc. of American Control Conference*, 2011, pp. 2578–2583.

[77] M. W. Spong, "Modeling and control of elastic joint robots," *Journal of Dynamic Systems, Measurement, and Control*, vol. 109, no. 4, pp. 310–319, 1987.

[78] J. Spooner and K. Passino, "Decentralized adaptive control of nonlinear systems using radial basis neural networks," *IEEE Transactions on Automatic Control*, vol. 44, no. 11, pp. 2050–2057, 1999.

[79] S. Stankovic, N. Ilic, Z. Djurovic, M. Stankovic, and K. Johansson, "Consensus based overlapping decentralized fault detection and isolation," in *Conference on Control and Fault-Tolerant Systems (SysTol'10)*, 2010, pp. 570–575.

[80] H. a. Talebi, K. Khorasani, and S. Tafazoli, "A recurrent neural-network-based sensor and actuator fault detection and isolation for nonlinear systems with application to the satellite's attitude control subsystem." *IEEE Transactions on Neural Networks*, vol. 20, no. 1, pp. 45–60, Jan. 2009.

[81] F. Thau, "Observing the state of non-linear dynamic systems," *International Journal of Control*, vol. 17, no. 3, pp. 471–479, 1973.

[82] B. Thumati and G. Halligan, "A Novel Fault Diagnostics and Prediction Scheme Using a Nonlinear Observer With Artificial Immune System as an Online Approximator," *IEEE Transactions on Control Systems Technology*, vol. 21, no. 3, pp. 569–578, 2013.

[83] A. Vemuri and M. Polycarpou, "Robust nonlinear fault diagnosis in input-output systems," *International Journal of Control*, vol. 68, no. 2, pp. 343–360, 1997. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/002071797223659

[84] V. Venkatasubramanian, R. Rengaswamy, S. Kavuri, and K. Yin, "A review of process fault detection and diagnosis:: Part III: Process history based methods," *Computers & Chemical Engineering*, vol. 27, no. 3, pp. 327–346, Mar. 2003.

[85] V. Venkatasubramanian, R. Rengaswamy, K. Yin, and S. Kavuri, "A review of process fault detection and diagnosis Part I: Quantitative model-based methods," *Computers & Chemical Engineering*, vol. 27, no. 3, pp. 293–311, Mar. 2003.

[86] V. Venkatasubramanian, R. Rengaswamy, and S. Kavuri, "A review of process fault detection and diagnosis:: Part II: Qualitative models and search strategies," *Computers & Chemical Engineering*, vol. 27, no. 3, pp. 313–326, 2003.

[87] L. Wei, W. Gui, Y. Xie, and S. X. Ding, "Decentralized Fault Detection System Design for Large-Scale Interconnected Systems," in *7th IFAC symposium on Fault Detection, Supervision and Safety of Technical Processes, Barcelona, Spain*, 2009, pp. 816–821.

[88] S. Xu, "Robust H_∞ filtering for a class of discrete-time uncertain nonlinear systems with state delay," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 49, no. 12, pp. 1853–1859, Dec. 2002.

[89] H. Yang and M. Saif, "Nonlinear adaptive observer design for fault detection," in *American Control Conference*, vol. 2, no. 2, 1995, pp. 1136–1139.

[90] ——, "State observation, failure detection and isolation (FDI) in bilinear systems," in *IEEE Conference on Decision and Control*, vol. 67, no. 6, Aug. 1997, pp. 2391–2396.

Christodoulos Keliris

[91] D. Yu, J. Gomm, D. Shields, D. Williams, and K. Disdell, "Fault diagnosis for, a gas-fired furnace using bilinear observer method," in *American Control Conference*, 1995, pp. 1127–1131.

[92] D. Yu and D. Shields, "A bilinear fault detection filter," *International Journal of Control*, vol. 68, no. 3, pp. 417–430, Oct. 1997.

[93] Q. Zhang and X. Zhang, "A distributed detection scheme for process faults and sensor faults in a class of interconnected nonlinear uncertain systems," in *IEEE 51st Conference on Decision and Control*, 2012, pp. 586–591.

[94] Q. Zhang, M. Basseville, and A. Benveniste, "Fault detection and isolation in nonlinear dynamic systems: A combined input-output and local approach," *Automatica*, vol. 34, no. 11, pp. 1359–1373, 1998.

[95] X. Zhang, T. Parisini, and M. M. Polycarpou, "Sensor bias fault isolation in a class of nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 50, no. 3, pp. 370–376, Mar. 2005.

[96] X. Zhang, M. M. Polycarpou, and T. Parisini, "A robust detection and isolation scheme for abrupt and incipient faults in nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 47, no. 4, pp. 576–593, 2002.

[97] ——, "Design and analysis of a fault isolation scheme for a class of uncertain nonlinear systems," *Annual Reviews in Control*, vol. 32, no. 1, pp. 107–121, 2008.

[98] ——, "Decentralized fault detection for a class of large-scale nonlinear uncertain systems," in *48h IEEE Conference on Decision and Control and 28th Chinese Control Conference*, 2009, pp. 6988–6993.

[99] X. Zhang, Q. Zhang, and N. Sonti, "Diagnosis of process faults and sensor faults in a class of nonlinear uncertain systems," *Journal of Systems Engineering and Electronics*, vol. 22, no. 1, pp. 22–32, 2011.

[100] Z. Zhang and I. Jaimoukha, "An optimal solution to an H-/H-infinity fault detection problem," in *In Proc. of 50th IEEE Conference on Decision and Control and European Control Conference*, 2011, pp. 903–908.

[101] Z. Zhang and I. M. Jaimoukha, "On-line fault detection and isolation for linear discrete-time uncertain systems," *Automatica*, vol. 50, no. 2, pp. 513–518, Feb. 2014. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0005109813005207

[102] F. Zhu and Z. Han, "A note on observers for Lipschitz nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 47, no. 10, pp. 1751–1754, 2002.

[103] P. J. Zufiria, "A formulation for fault detection in stochastic continuous-time dynamical systems," *International Journal of Computer Mathematics*, vol. 86, no. 10, pp. 1778–1797, Oct. 2009.

Christoforos Keliris