Fall 2013

Model based fault diagnosis and prognosis of nonlinear systems

Hasan Ferdowsi

Department: Electrical and Computer Engineering

MODEL BASED FAULT DIAGNOSIS AND PROGNOSIS OF NONLINEAR SYSTEMS

by

HASAN FERDOWSI

A DISSERTATION

Presented to the Faculty of the Graduate School of the

MISSOURI UNIVERSITY OF SCIENCE AND TECHNOLOGY

In Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

in

ELECTRICAL ENGINEERING

2013

Approved by

Jagannathan Sarangapani, Advisor
Levent Acar
S.N. Balakrishnan
Cihan Dagli
Maciej Zawodniok
Al Salour

PUBLICATION DISSERTATION OPTION

This dissertation contains the following five articles:

Paper 1: Hasan Ferdowsi and S. Jagannathan, "A Unified Model-Based Fault Diagnosis Scheme for Nonlinear Discrete-Time Systems with Additive and Multiplicative Faults," in *Transactions of the Institute of Measurement and Control,* Vol. 35, No. 6, pp. 742-752, 2013.

Paper 2: Hasan Ferdowsi, S. Jagannathan, and M. Zawodniok, "An Online Outlier Identification and Removal Scheme for Improving Fault Detection Performance," accepted for publication in *IEEE Transactions on Neural Networks.*

Paper 3: Hasan Ferdowsi and S. Jagannathan, "Decentralized Fault Diagnosis and Prognosis Scheme for Interconnected Nonlinear Discrete-Time Systems," revised and under review with *IEEE Transaction on Neural Networks*.

Paper 4: Hasan Ferdowsi and S. Jagannathan, "A Decentralized Fault Detection and Accommodation Scheme for Interconnected Nonlinear Continuous-time Systems," under review with *IEEE Transactions on Systems, Man, and Cybernetics.*

Paper 5: Hasan Ferdowsi and S. Jagannathan, "Fault Diagnosis of a Class of Distributed Parameter Systems Modeled by Parabolic Partial Differential Equations," To be submitted to *Automatica.*

ABSTRACT

Rapid technological advances have led to more and more complex industrial systems with significantly higher risk of failures. Therefore, in this dissertation, a model-based fault diagnosis and prognosis framework has been developed for fast and reliable detection of faults and prediction of failures in nonlinear systems.

In the first paper, a unified model-based fault diagnosis scheme capable of detecting both additive system faults and multiplicative actuator faults, as well as approximating the fault dynamics, performing fault type determination and time-to-failure determination, is designed. Stability of the observer and online approximator is guaranteed via an adaptive update law. Since outliers can degrade the performance of fault diagnostics, the second paper introduces an online neural network (NN) based outlier identification and removal scheme which is then combined with a fault detection scheme to enhance its performance. Outliers are detected based on the estimation error and a novel tuning law prevents the NN weights from being affected by outliers.

In the third paper, in contrast to papers I and II, fault diagnosis of large-scale interconnected systems is investigated. A decentralized fault prognosis scheme is developed for such systems by using a network of local fault detectors (LFD) where each LFD only requires the local measurements. The online approximators in each LFD learn the unknown interconnection functions and the fault dynamics. Derivation of robust detection thresholds and detectability conditions are also included. The fourth paper extends the decentralized fault detection from paper III and develops an accommodation scheme for nonlinear continuous-time systems. By using both detection and accommodation online approximators, the control inputs are adjusted in order to minimize the fault effects.

Finally in the fifth paper, the model-based fault diagnosis of distributed parameter systems (DPS) with parabolic PDE representation in continuous-time is discussed where a PDE-based observer is designed to perform fault detection as well as estimating the unavailable system states. An adaptive online approximator is incorporated in the observer to identify unknown fault parameters. Adaptive update law guarantees the convergence of estimations and allows determination of remaining useful life.

ACKNOWLEDGEMENTS

It would not have been possible to complete this dissertation without the help and support of the kind people around me, to only some of whom it is possible to give particular mention here.

I would like to express my deepest gratitude to my advisor, Professor Jagannathan Sarangapani for his excellent guidance, support, and patience over the duration of my doctoral research endeavor. I have been extremely lucky to have an advisor who cared so much about my work. His recommendations and consistent encouragement enabled me to effectively perform the research and complete this dissertation. I am also grateful to Dr. Maciej Zawodniok for his advice and contributions to my doctoral research projects. Furthermore, I would like to thank the rest of my committee members, Dr. L. Acar, Dr. S. N. Balakrishnan, Dr. C. Dagli, and Dr. A. Salour, for their valuable guidance. In addition, I acknowledge the financial support provided by the National Science Foundation.

I must extend my appreciation to my wife without whom this effort would have been worth nothing. She has supported me in every possible way to see the completion of this dissertation. I am very much indebted to my parents and two brothers, who encouraged and helped me at every stage of my personal and academic life and I am eternally grateful to them.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

1. INTRODUCTION

Faults and failures are inevitable in any kind of industrial systems. The risk of failure is high with an increase in system complexity. Therefore, reliable fault diagnosis and prognosis schemes are required to guarantee the safety of system operation and human operators, and to minimize the risk of irreversible damage to components.

In the past decade, significant advances in theoretical and applied research have occurred in the area of fault diagnosis. The fault diagnosis schemes are based on either data-driven or model-based. Data-driven methods [1] can be very useful when the mathematical model of a linear or nonlinear system is not available. What is common among all data-driven methods is the need for data from both healthy and fault operating conditions of the system under consideration. Therefore, it is more difficult to design a generic data-driven fault diagnosis method applicable to a wide range of systems. Moreover, collecting measurements in faulty conditions can be very costly and in some cases even impossible.

In contrast, model-based methods [2] minimize the need for a priori data and can perform online, but they require accurate mathematical model of the system. However, the two aforementioned classes of fault diagnosis have become closer, as researchers have recently been trying to combine both methods, in order to eliminate the disadvantages of each method and construct more reliable and functional fault diagnosis schemes [3,4]. For instance data, if available, can be utilized to tune the system model and also to determine robust fault detection thresholds. In both the cases, fault diagnosis is done whereas prognosis is still in its infancy.

The design of all model-based fault diagnosis schemes starts with development of an observer, based on the available system model [5]. Figure 1.1 illustrates the general structure of model-based fault detection. The observer provides estimates of system states in healthy operating conditions though the system state vector is measured. A residual signal is generated by comparing the measured and estimated system states. As long as the system is working under healthy operating conditions, the difference between actual and estimated states will be less than a certain threshold, provided the observer is designed appropriately. When a fault occurs, the actual system dynamics will be changed,

but the observer dynamics remains unaltered. As a result, the actual system states will deviate from the estimated states, thus increasing the detection residual. Fault detection is performed by comparing the residual with a predefined detection threshold.

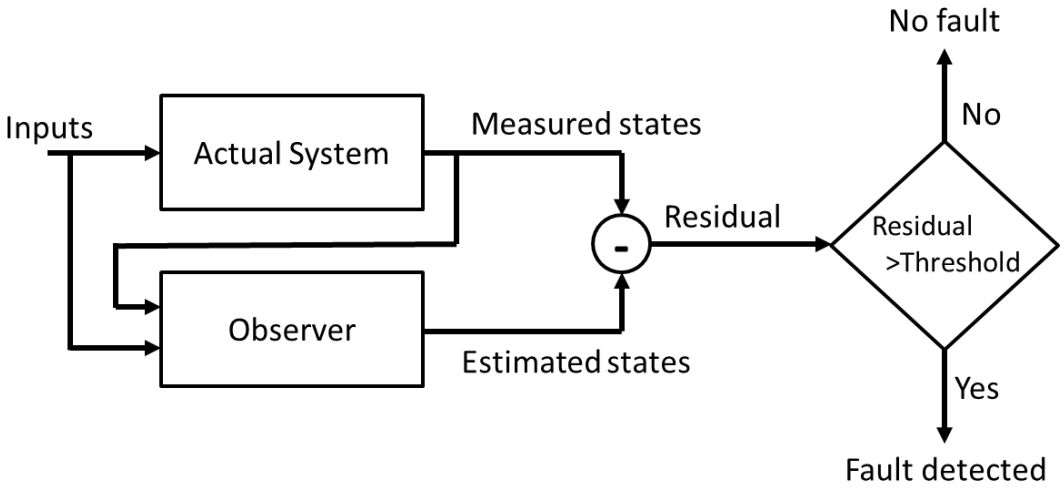

Figure 1.1. Model-based fault detection

Although the detection of a fault is the most important step, it is only a first step. When a fault is detected, the estimation of fault dynamics, determination of fault type and its location become important, since they can facilitate the root cause analysis of faults for system maintenance and repair. Moreover, it is imperative to determine the remaining life of the system in order to improve the system availability and prevent either component damage or complete system failure.

Given the importance of fault diagnosis and prognosis, this topic has attracted a large number of researchers who have worked on different aspects of it over the past decade. An overview of current fault diagnosis methodologies and their shortcomings are discussed next. Subsequently, the contributions of this dissertation are introduced.

1.1. OVERVIEW OF FAULT DIAGNOSIS METHODOLOGIES

Many different approaches to model-based fault detection and diagnosis has been introduced recently, all of which require an estimator/observer. Fault diagnosis is performed based on adaptive estimators in [6], by using neural network (NN) based estimators in [7], and by utilizing fuzzy observers in [8].

Based on their mathematical representation, faults can be classified into two classes of additive and multiplicative faults. Additive fault representation, which is generally used to model system and component faults, is very common in the fault detection literature including [9,10] and previously mentioned model-based FD literature [6-8]. On the other hand, actuator faults which result in partial loss of control action are commonly modeled as multiplicative faults.

Fault diagnosis of systems with multiplicative actuator faults have been done by utilizing parameter similarity measures [11], sliding mode observers [12], and NN based techniques [13]. However, all of these FD methods [6-13] are designed for continuous time systems, and more importantly each of them can only handle either additive or multiplicative faults, whereas practical systems can be subjected to both types of fault and the fault diagnosis should be capable of detecting them and determination of detected fault type.

Another need of the current fault diagnosis schemes [6-13], is that they do not consider noise and outliers in measurements. An outlier, by definition, is an observation which deviates significantly from other observations thus creating suspicion that it was generated by a different system. The measured data in industrial systems usually involve noise and outliers, which not only degrades the data quality but also can render inaccurate decisions during fault diagnosis. Therefore, reliable fault detection schemes are required to perform the fault detection online without missed or false alarms due to outliers. For this purpose, preprocessing of measured data is necessary to detect and remove outliers before they can affect the fault detection decision.

Several outlier detection schemes have been proposed in the literature, such as distribution-based [14], distance-based [15], and density based methods [16]. However, these methods cannot work online, thus impossible to be implemented as a preprocessing unit for online fault detection. Therefore, online outlier detection and removal methods

have also been proposed recently by using Kalman filter and its variations [17,18]. These methods need the system dynamics to be known and fixed and assume that the underlying distribution of system states is fixed, whereas this assumption cannot always be satisfied in practical systems due to change in system operating conditions. Moreover, when a fault occurs in the system, it can change both the system dynamics and the distribution of states. Therefore, a novel scheme is required for online detection and removal of outliers for nonlinear systems in nonstationary conditions.

Another important problem in the area of fault diagnosis is related to the decentralized systems. Several practical systems such as power generation and distribution systems, telecommunication networks, traffic networks, etc, exhibit complex and spatially distributed dynamics and are referred to as large scale interconnected systems. The aforementioned FD literature [6-13] addresses centralized schemes which are not suitable for distributed systems. Due to the extensive effort required in transmitting the entire system measurements for a centralized scheme, decentralized control of distributed systems by using local subsystem states is introduced recently [19,20]. Distributed fault diagnosis schemes have also been proposed for dealing with large scale interconnected systems assuming that the entire system states or entire estimated states are available at all subsystems. Since it is very expensive and time consuming to gather and process all the measurements from a distributed large scale system at one place and the measurements can be outdated due to delay in transmission, the need for a pure decentralized FD scheme which only uses local measurements at each subsystem is desirable. Moreover, the remaining useful life information is not included in the above schemes [19,20].

Although certain faults are critical and the overall system must be forced to shut down upon their detection, other faults at an incipient stage can be accommodated for a limited time, which allows uninterrupted operation of the system with a desired performance in the presence of faults. Many centralized fault accommodation schemes like observer-based methods [21,22] have been proposed. As previously mentioned, large scale interconnected systems require decentralized schemes, which motivated researchers to work on distributed approaches for fault accommodation [23,24]. Similar to current distributed fault diagnosis, these accommodation methods are not completely

decentralized and still require the interconnection functions to be known and the entire state vector to be available at all subsystems. In contrast, this dissertation proposes decentralized fault diagnosis and fault accommodation schemes by using only the local states at each local fault detector.

Another class of systems which require careful attention is distributed parameter systems (DPS). In such systems, the variables evolve both in time and space in contrast with other nonlinear systems mentioned above that evolve only with time. Distributed parameter systems are generally described by partial differential equations (PDE) in contrast to ordinary lumped parameter systems which are described by ordinary differential equations (ODE) [25]. Current DPS fault diagnosis methods, like [26-28] require the transformation of PDE model to a finite dimensional ODE using Galerkin's method [29,30], which can render inaccurate results due to certain physical aspects of the system being neglected. Furthermore, when a fault happens in the system, the system PDE dynamics will change, which can make the approximated ODE model even more inaccurate. This is the motivation behind the last chapter of this dissertation, which is the design of a more accurate and reliable fault diagnosis and prognosis method for distributed parameter systems.

1.2. ORGANIZATION OF THE DISSERTATION

In this dissertation, novel model-based fault diagnosis and prognosis schemes are introduced for classes of nonlinear systems which are experienced by faults. In addition, a new online outlier identification and removal scheme which can be used as a preprocessing unit for fault detection is introduced. The dissertation, which is illustrated in Figure 1.2, is presented in the form of five papers. The common theme in the five papers is the model-based fault prognosis of nonlinear systems. The first three papers deal with discrete-time systems, whereas the fourth and fifth papers present methods for continuous-time systems.

In the first paper, a unified model-based fault diagnosis scheme for nonlinear systems which is capable of detecting both additive system faults and multiplicative actuator faults. Fault is detected using the detection residual generated by comparing the actual and estimated states. Two online approximators, which are both offline prior to

fault detection, are incorporated in the nonlinear observer. Upon detection, the first online approximator is turned on to learn the input signal. Subsequently an input residual is generated and compared against a user-defined threshold to identify the type of fault. Identification of fault type allows the activation of appropriate online approximator to learn the fault dynamics. Lyapunov techniques are used to show that detection residual and parameter estimation errors are uniformly ultimately bounded. Time-to-failure is also determined by using the parameter update law, the active online approximator parameters, and their designer specified failure thresholds.

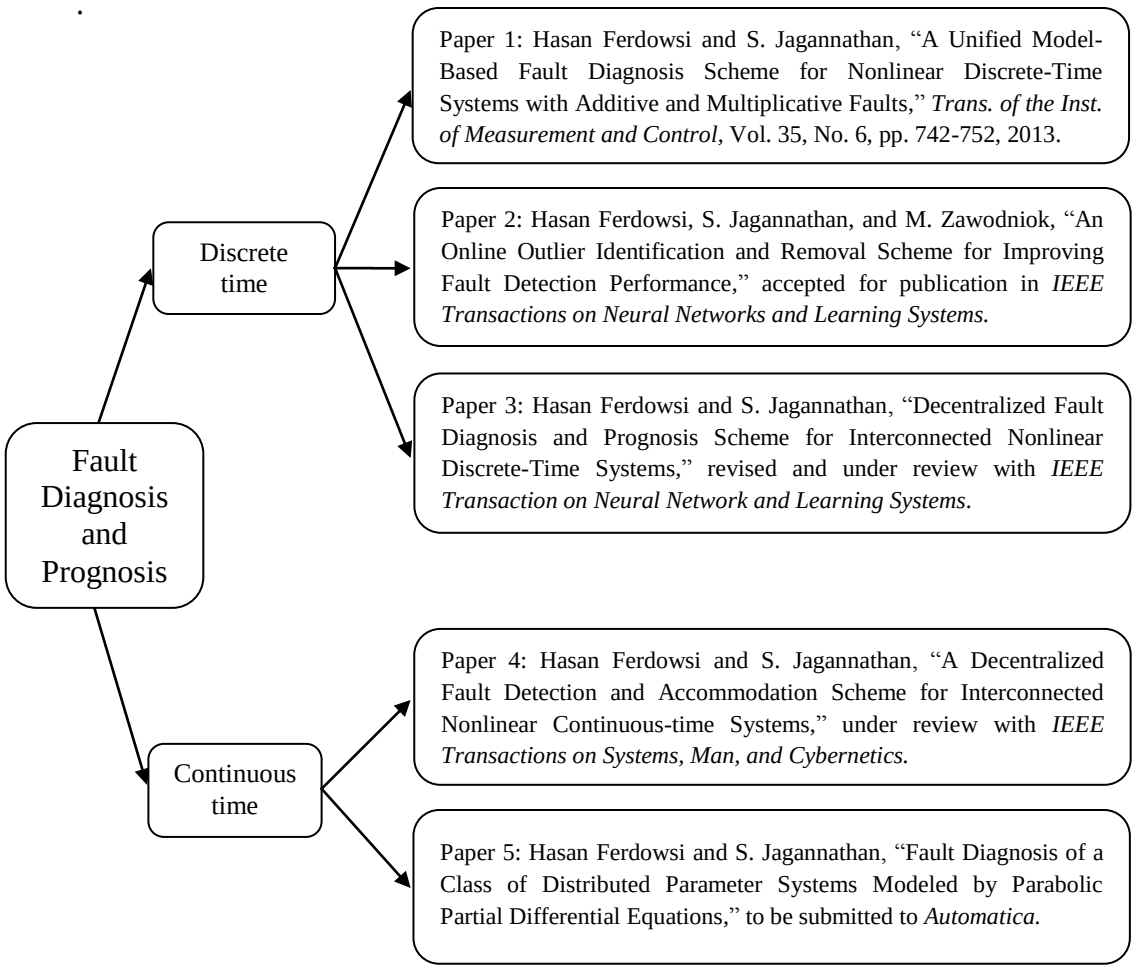

Figure 1.2. Dissertation outline

The second paper deals with the problem of outliers in measured data. An online NN based method of outlier identification and removal (OIR), which acts as a prefilter to detect and eliminate outliers from measured data before fault detection, is introduced. A two layer feedforward NN is used to estimate the system states and state estimation error is defined as the difference between measured and estimated states. At each time instant, median and standard deviation of the state estimation error is calculated over a limited time window. Then, an outlier is detected if the difference between state estimation error and the calculated median is higher than three standard deviations. A novel adaptive weight update law is used to prevent the update of NN weights by the detected outlier. For the purpose of fault detection, a different observer with known nominal dynamics of the system is introduced. This observer uses the estimated outlier-free states instead of the measured states which contain outliers. The stability of both the OIR and the fault detection scheme are discussed in the paper.

In the third paper, a new decentralized fault diagnosis and prognosis scheme for discrete-time nonlinear systems is presented. A network of local fault detectors (LFD) is designed, where each LFD only requires the local system states and inputs. Each local fault detector is constructed as a nonlinear observer which includes an online approximator. A residual is generated by comparing the estimated and measured states, but it is only used to update the unknown parameters of online approximator. The online approximator is activated all the time, and is used to approximate the interconnection function and the possible fault function. Since the magnitude of interconnection term is supposed to be bounded in healthy operating conditions, the detection of fault is performed by comparing the output of online approximator with the bound on interconnection term. Robust detection thresholds, fault detectability conditions, and time-to-failure determination formula are mathematically derived and a fault isolation algorithm is also included to determine the location of fault.

Subsequently, the decentralized fault detection and accommodation of a class of large scale interconnected systems is addressed in the fourth paper. Each local fault detector involves an online approximator (OLA) which is turned off prior to detection. When the detection residual in one subsystem crosses the detection threshold, a fault is detected and the online approximator in that subsystem is activated to learn the fault

dynamics. At the same time, the online approximators in all the other subsystems are notified to remain offline. Upon detection, an accommodation online approximator is also initiated at each subsystem. The outputs of detection OLA along with the accommodation OLA are utilized to modify the control inputs to minimize the effect of fault. Lyapunov proofs are offered for the local fault detection and accommodation schemes. Moreover, time-to-failure and time-to-accommodation are introduced, which allow determination of whether the accommodation unit can bring the system performance back to normal before the system reaches a failure.

Finally, in the last chapter, the fault diagnosis of distributed parameter systems is investigated. A fault detection observer is designed by directly utilizing the nominal PDE model of the system instead of the approximated ODE representation. Residual is generated for fault detection by comparing estimated and actual system outputs. An adaptive approximator which is incorporated in the observer is activated upon detection in order to identify unknown fault parameters and approximate fault dynamics. Adaptive parameter update law guarantees the observer convergence and allows time-to-failure determination.

1.3. CONTRIBUTIONS OF THE DISSERTATION

The contributions of the first paper include the development of a unified fault diagnosis and prognosis scheme via a novel observer for detecting both multiplicative actuator and additive system faults in contrast with the literature [6-13] where a single fault type is normally handled. Fault type is identified by using the input residual and fault detectability conditions and upper bound for detection time are derived analytically depending upon the fault type. Subsequently an online TTF determination scheme is introduced while such schemes are unavailable in the literature for model-based methods.

The contributions of second paper involve the design of a new outlier detection and removal scheme, which can operate online in contrast with data-driven methods [14-16]. This scheme can be applied to both linear and nonlinear systems in nonstationary environments in compaison with existing Kalman filter-based schemes [17,18]. The proposed OIR scheme is combined with a model-based fault detection scheme, to enhance its performance in the presence of noise and outliers. However, the OIR scheme

can be utilized both for data driven and model-based fault diagnosis schemes, since it is generic and does not use the system representation or model.

The major contributions of the third paper include the development of a decentralized fault diagnosis scheme for nonlinear discrete-time systems wherein a LFD only uses local measurements in contrast with [19,20]. Each LFD is designed to detect faults regardless of their location and then determine if the fault is local or nonlocal by using a centralized isolation module. Furthermore TTF estimation is performed upon fault detection whereas such scheme is not available in most of the model-based methods [6-13,19,20].

On the other hand, the fourth paper contributes to the field of fault diagnosis and prognosis by proposing a new decentralized fault detection and accommodation scheme for nonlinear interconnected continuous-time systems. Both the detection and accommodation schemes use only local subsystem states and inputs to detect and accommodate the faults in contrast with existing methods [21-24]. In addition, analytical formulas for online calculation of time-to-failure and time-to-accommodation are derived. Based on the information, the human operator can make a decision on keeping the system running or shutting it down based on safety assessment.

Finally the fifth paper considers the problem of fault diagnosis and prognosis of distributed parameter systems represented as parabolic PDE models. In contrast with existing schemes [26-28], the PDE model of the system is not going to be transformed into a finite dimensional ODE model before performing the fault detection. Instead, the detection observer is going to be designed based on the PDE model with an incorporated online approximator. This method allows more accurate estimation of states, thus providing more reliable fault detection and approximation results.

1.4. REFERENCES

[1] S. M. Namburu, M. Wilcutts, S. Chigusa, Q. Liu, C. Kihoon, and K. Pattipati, "Systematic data-driven approach to real-time fault detection and diagnosis in automotive engines," *IEEE Autotestcon,* pp. 59-65, 2006.

[2] R. Isermann, "Model-based fault-detection and diagnosis – status and applications," *Annual Reviews in Control,* vol. 29, pp. 71-85, 2005.

[3] B. J. Ohran, J. Liu, D. M. de la Peña, P. D. Christofides, and J. F. Davis, "Data-based fault detection and isolation using feedback control: Output feedback and optimality," *Chemical Engineering Science,* vol. 64, pp. 2370-2383, 2009.

[4] R. Ghimire, C. Sankavaram, A. Ghahari, K. Pattipati, Y. Ghoneim, M. Howell, and M. Salman, "Integrated model-based and data-driven fault detection and diagnosis approach for an automotive electric power steering system," *IEEE Autotestcon*, pp. 70-77, 2011.

[5] J. 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.

[6] M. A. Demetriou and M. M. Polycarpou, "Incipient fault diagnosis of dynamical systems using online approximators," *IEEE Trans. on Automatic Control*, vol. 43, no. 11, pp. 1612-1617, 1998.

[7] A. Alessandri and T. Parisini, "Model-based fault diagnosis using nonlinear estimators: a neural approach," *Proceedings of the 1997 American Control Conference*, vol. 2, pp. 903-907, 1997.

[8] C. J. Lopez-Toribio and R. J. Patton, "Fuzzy observers for nonlinear dynamic systems fault diagnosis," *Proc. IEEE Conf. on Decision and Control*, pp. 84–89, 1998.

[9] A. Alessandri, "Fault diagnosis for nonlinear systems using a bank of neural estimators," *Comp. in Industry*, vol. 52, no. 3, pp. 271-289, 2003.

[10] F. Caccavale and L. Villani, "An adaptive observer for fault diagnosis in nonlinear discrete-time systems," *Proc. of the American Control Conference,* vol. 3, pp. 2463-2468, 2004.

[11] H. P. Huang, C. C. Li, and J. C. Jeng, "Multiple multiplicative fault diagnosis for dynamic processes via parameter similarity measures," *Industrial & Engineering Chemistry Research*, vol. 46, no. 13, pp. 4517-4530, 2007.

[12] T. Chee Pin and C. Edwards, "Multiplicative fault reconstruction using sliding mode observers," *5th Asian Control Conference*, vol. 2, pp. 957-962, 2004.

[13] H.A. Talebi, and K. Khorasani, "A neural network-based actuator gain fault detection and isolation strategy for nonlinear systems," *IEEE Conference on Decision and Control*, vol. 46, pp. 2614-2619, 2007

[14] P. J. Rousseeuw and K. V. Driessen, "A fast algorithm for the minimum covariance determinant estimator," *Technometrics,* vol. 41, pp. 212-223, 1999.

[15] E. M. Knorr, R. T. Ng, and V. Tucakov, "Distance-based outliers: algorithms and applications," *The VLDB Journal*, vol. 8, pp. 237-253, 2000.

[16] Z. Jiang-She and L. Yiu-Wing, "Robust clustering by pruning outliers," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 33, pp. 983-998, 2003.

[17] G. Agamennoni, J. I. Nieto, and E. M. Nebot, "An outlier-robust Kalman filter," *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1551-1558, 2011.

[18] S. C. Chan, Z. G. Zhang, and K. W. Tse, "A new robust Kalman filter algorithm under outliers and system uncertainties," *IEEE International Symposium on Circuits and Systems*, vol. 5, pp. 4317-4320, 2005.

[19] R. M. G. Ferrari, T. Parisini, and M. M. Polycarpou, "Distributed fault diagnosis with overlapping decompositions: an adaptive approximation approach," *IEEE Trans on Automatic Control*, vol. 54, no. 4, pp. 794–799, 2009.

[20] S. Stankovic, N. Ilic, Z. Djurovic, M. Stankovic, K.H. Johansson, "Consensus based overlapping decentralized fault detection and isolation," *Proc. 2010 Conference on Control and Fault-Tolerant Systems (SysTol'10)*, pp. 570–575, 2010.

[21] S. N. Huang and T. Kok Kiang, "Fault detection, isolation, and accommodation control in robotic systems," *IEEE Transactions on Automation Science and Engineering*, vol. 5, pp. 480-489, 2008.

[22] J. Bin, M. Staroswiecki, and V. Cocquempot, "Fault accommodation for nonlinear dynamic systems," *IEEE Transactions on Automatic Control*, vol. 51, pp. 1578-1583, 2006.

[23] P. Panagi and M. M. Polycarpou, "Distributed fault accommodation for a class of interconnected nonlinear systems with partial communication," *IEEE Transactions on Automatic Control*, vol. 56, pp. 2962-2967, 2011.

[24] J. D. Boskovic and R. K. Mehra, "A decentralized scheme for accommodation of multiple simultaneous actuator failures," *Proceedings of the 2002 American Control Conference.*, vol. 6, pp. 5098-5103, 2002.

[25] S. Omatu and J. H. Seinfeld, *Distributed parameter systems: theory and applications*. Clarendon Press, 1989.

[26] M. A. Demetriou, "A model-based fault detection and diagnosis scheme for distributed parameter systems: a learning systems approach," *ESAIM: Control, Optimisation and Calculus of Variations*, vol. 7, pp. 43-67, 2002.

[27] S. Ghantasala and N.H. El-Farra, "Detection, isolation and management of actuator faults in parabolic PDEs under uncertainty and constraints," *46th IEEE Conference on Decision and Control*, pp. 878-884, 2007.

[28] A. Baniamerian and K. Khorasani, "Fault detection and isolation of dissipative parabolic PDEs: Finite-dimensional geometric approach," *American Control Conference (ACC)*, pp. 5894-5899, 2012.

[29] P. D. Christofides, *Nonlinear and Robust Control of PDE Systems: Methods and Applications to transport reaction processes*. Springer, 2001.

[30] A. Friedman, *Partial Differential Equations*. Dover Publications, Incorporated, 2008.

PAPER

I. A Unified Model-Based Fault Diagnosis Scheme for Nonlinear Discrete-Time Systems with Additive and Multiplicative Faults

Hasan Ferdowsi and S. Jagannathan

Abstract

In this paper, a unified model-based fault diagnosis (MFD) scheme that deals with both multiplicative actuator and additive system faults is designed. For a class of uncertain nonlinear discrete-time systems, this MFD scheme is capable of not only detecting both additive and multiplicative actuator faults but also to identify the fault type. Faults are detected by using a novel fault detection observer (FD) consisting of two online approximators in discrete-time (OLAD) and a robust adaptive term. Upon detection, a fault diagnosis scheme is introduced to determine the fault type by monitoring the input residual generated via the first OLAD output. Upon performing the diagnosis online, the appropriate OLAD is activated in the observer and the other OLAD is switched off. Thereafter, by using both the parameter update law of the active OLAD and user-selected failure threshold, an online time-to-failure (TTF) scheme is introduced. In the case of multiplicative faults, boundedness of the detection residual and parameter estimation errors is shown while in the case of additive faults, the asymptotic convergence of the detection residual and parameter estimation errors is guaranteed due to the robust adaptive term. Finally a simulation example is used to demonstrate the proposed fault diagnosis scheme.

1. INTRODUCTION

Today's industrial systems have become more complex and are prone to failures. Faults are considered as a precursor to failures and they can occur both in the sensors and actuators. Therefore, reliable fault diagnosis schemes are required to guarantee safe operation of the physical system even in the presence of uncertainties and faults. If the faults can be detected early enough, performance loss and damage to the system can be prevented.

Faults are classified based on their representation into additive and multiplicative faults. Additive fault representation, which is very popular in the fault diagnosis literature, is used to model system and component faults, whereas actuator faults which represent partial loss of control action are commonly modeled as multiplicative faults. Despite the fault type, the diagnosis schemes are generally divided into data-driven and model-based methods. Data-driven fault diagnosis approaches need healthy and faulty data [1] for each fault which can be very expensive. In addition, an offline training session is normally needed. As a result, these methods are not preferred since they result in false alarms when the operating conditions change.

On the other hand, model-based fault diagnosis methods [2] minimize the need for a priori data and can detect faults online. In this approach, an observer or estimator representative of the system is utilized for detecting faults [2-4]. The observer provides an estimate of the system states. Then a residual signal is generated by comparing the estimated states with that of the actual system states, and when this residual exceeds a predefined threshold, a fault is detected [3]. These model-based FD methods have been implemented on both linear and nonlinear systems that have a linear representation [5,6]. Availability of a priori data during healthy and faulty operation from the system can aid in the fine tuning of the model-based schemes.

As part of model-based FD framework, in [7,8], fault diagnosis schemes using adaptive estimators have been discussed while in [9,10] neural network (NN)-based estimators have been utilized for the purpose of fault detection. Fuzzy observers have been utilized in certain model-based fault detection schemes [11,12].

Numerous researchers have worked on the detection of additive faults [7,8,10,13,14]. On the other hand, fault diagnosis for systems with multiplicative faults

has also been done using a fuzzy observer [11], parameter similarity measures[15], sliding mode observer [16], and NN-based methods [17]. However all these papers [11,15-17] deal with continuous time systems, whereas in this paper discrete time systems are considered. Moreover in this paper, a fault diagnosis scheme, which is designed for multiplicative faults, enables detection threshold selection in an analytical manner, and that is utilized to determine time-to-failure (TTF) upon detection, in contrast to the literature. Detectability conditions are also introduced. Further, in all these papers [7,8,10-14,16,17] only one type of fault is normally considered while practical nonlinear systems can be subject to both additive and multiplicative faults which is the focus of this paper.

Therefore, the nonlinear discrete-time system considered in this paper is subjected to both additive and multiplicative faults by assuming the entire state vector is available for measurement. A novel observer design is proposed wherein two OLADs, one for each fault type, is introduced as part of the unified prognosis framework. A neural network (NN) is used as an OLAD [10] whereas any online approximator can be utilized. Initially, prior to detection, both the OLADs and robust adaptive term are zero. Detection residual is generated first by comparing the estimates state vector of the observer with that of the nonlinear system. A deadzone operator is used to declare the presence of a fault when the detection residual exceeds a user defined threshold.

Upon detection, the first OLAD is activated to learn the fault dynamics and to generate an estimate of the input signal using which an input residual is obtained. This input residual is compared against a user defined threshold to identify the type of fault that has occurred. Upon identifying the fault type, a decision is made to activate the appropriate OLAD. In other words, upon detection, only one OLAD out of the two will be active at any given time.

Next, TTF is determined online by using a mathematical equation relating to the active OLAD parameters. In this mathematical equation, the appropriate parameter estimates are compared against the designer specified limits since for most practical systems the parameters could be tied to physical quantities which have a safe range of values [13]. Then the overall TTF for the system can be found by taking the minimum TTF for all system parameters. In this paper, TTF determination is performed for the

system both with additive or multiplicative faults provided a single fault type can occur at a given time.

Thus the contributions of this paper include the development of a unified fault diagnosis and prognosis scheme via a novel observer for detecting both multiplicative actuator and additive system faults in contrast with the literature [7,13,15-17,20] where a single fault type is normally handled. The novel observer design includes two OLADs in contrast with traditional schemes [2,5,8,9,13] where one OLAD is utilized. Fault type is identified by using the input residual and then the appropriate OLAD is used to estimate additive or multiplicative fault dynamics. Fault detectability conditions and upper bound for detection time are derived analytically depending upon the fault type. Subsequently an online TTF determination scheme is then introduced while such schemes are unavailable in the literature for model-based methods [7,11,14,20]. In this framework, a model-based additive fault detection scheme from [13] is utilized for demonstrating the unified framework.

This paper is organized as follows: Section 2 describes the system under consideration and the possible faults while Section 3 discusses the FD and diagnosis scheme to identify the type of fault. Section 4 describes the TTF determination. In Section 5, a simulation example is used to verify the performance of the proposed scheme and finally Section 6 presents conclusions and future works.

2. SYSTEM DESCRIPTION

Consider the nonlinear discrete-time system described by

$$x\left(k+1\right)=\omega\left(x\left(k\right),u\left(k\right)\right)+\eta\left(x\left(k\right),u\left(k\right)\right)$$

where $u\in\mathbb{R}^m$ is the control input vector, $x\in\mathbb{R}^n$ is the system state vector which is assumed to be measurable, $\omega:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$ represents the known nonlinear system dynamics, and $\eta:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$ represents the system uncertainties. Due to an additive fault, the nonlinear system is given by

$$x\left(k+1\right)=\omega\left(x\left(k\right),u\left(k\right)\right)+\eta\left(x\left(k\right),u\left(k\right)\right)+\Pi\left(k-k_0\right)h\left(x\left(k\right),u\left(k\right)\right)$$

where $h(x(k),u(k))$ represents a vector of possible additive fault dynamics, which are defined as $h(.)=\left[\theta_1^T f_1\left(x(k),u(k)\right),\ldots,\theta_n^T f_n\left(x(k),u(k)\right)\right]$. $\theta_i \in \mathbb{R}^{l_i}$, $i=1,\ldots,n$, is an unknown parameter vector, and $f_i : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^{l_i}$,$i=1,\ldots,n$, is a known fault basis function. Each f_i represents the fault function of the ith fault affecting the ith state equation, and each θ_i is the unknown magnitude of the ith fault function. The time profile of a fault is given by $\Pi(k-k_0)$.

Let $T_A\left(x(k),u(k)\right):\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ denote the vector of unknown multiplicative actuator faults [17], and therefore $u_f(k)=diag\left\{T_A\left(x(k),u(k)\right)\right\}u(k)$. Now during the fault conditions, we replace $u(k)$ by $u_f(k)$, then the nonlinear system dynamics with the additive and multiplicative faults can be rewritten as

$$x(k+1)=\omega\left(x(k),u_f(k)\right)+\eta\left(x(k),u_f(k)\right)+\Pi(k-k_0)h\left(x(k),u_f(k)\right) \qquad (1)$$

The time profile $\Pi(k-k_0)$ is modeled by

$$\Pi(k-k_0)=diag\left\{\Omega_1(k-k_0),\Omega_2(k-k_0),\ldots,\Omega_n(k-k_0)\right\}$$

where

$$\Omega_i(k-k_0)=\begin{cases}0, & if\ k<k_0\\ 1-e^{-\bar{\kappa}_i(k-k_0)}, & if\ k\geq k_0\end{cases} \quad for\ i=1,\ldots,n$$

and $\bar{\kappa}_i$ is an unknown constant that represents the rate at which a fault occurs. A larger value of $\bar{\kappa}_i$ indicates that it is an abrupt fault. The use of such time profiles is common in fault diagnosis literature [7,13,18]. Next standard assumptions are needed in order to proceed.

Assumption 1: The modeling uncertainty is bounded, i.e. $\left\|\eta\left(x(k),u(k)\right)\right\|\leq\eta_M,\forall(x,u)\in(X,U)$, where η_M is a positive known constant and X and U define the range of possible states and inputs.

Assumption 2: Only a single fault type of either multiplicative or additive can occur in the system at any given time. This means that both additive and multiplicative

faults cannot happen simultaneously while multiple faults of the same type can still occur.

Remark 1: Assumption 1 is needed to distinguish between faults and system uncertainties while Assumption 2 is needed to identify the fault types.

Assumption 3: The nonlinear system dynamics $\omega(x,u)$ is Lipschitz in u, i.e., $\|\omega(x(k),u_1(k))-\omega(x(k),u_2(k))\| \le c_g \|u_1(k)-u_2(k)\|$, where $c_g > 0$ is the Lipschitz constant [20].

Assumption 4: For the purpose of TTF, fault functions can be expressed as nonlinear in the unknown parameters (NLIP) [19], i.e. both additive and multiplicative fault functions can be approximated by two-layer NN with bounded activation functions and weight parameters.

Assumption 5: Due to finite actuator bandwidth, the control input u has a finite upper bound such that $\|u(k)\| \le u_{max}$ where u_{max} is a positive constant.

For practical systems, actuator output is finite and the boundedness assumption is also considered in the fault diagnosis literature [7,8]. Next the proposed fault diagnosis scheme is introduced.

3. FAULT DIAGNOSIS SCHEME

In this section, the proposed fault diagnosis scheme for detecting additive and multiplicative faults will be described. For this purpose, we first introduce the fault detection (FD) estimator by which both types of faults can be detected. Upon detection of a fault, a method of identifying the fault type is introduced. Finally the stability of the detection residual is demonstrated for the nonlinear system subjected to faults.

Consider the nonlinear FD estimator or observer

$$\begin{aligned} \hat{x}(k+1) &= A_d \hat{x}(k) + \omega\left(x(k), \hat{u}_f\left(u(k), \hat{T}_A(k)\right)\right) + \hat{h}_d\left(x(k), u(k); \hat{\theta}_d(k)\right) - A_d x(k) + F(k) \\ \hat{u}_f\left(u(k), \hat{T}_A(k)\right) &= diag\left\{\hat{T}_A\left(x(k), u(k); \hat{\theta}_A(k)\right)\right\} u(k) \end{aligned} \tag{2}$$

where $\hat{x}(k) \in \mathbb{R}^n$ is the estimated state vector, $\hat{T}_A : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{q \times m} \to \mathbb{R}^m$ is the output of the first OLAD with $\hat{\theta}_A \in \mathbb{R}^{q \times m}$ being its set of adjustable parameters,

$\hat{h}_d : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{p\times n} \to \mathbb{R}^n$ is the output of the second OLAD with $\hat{\theta}_d \in \mathbb{R}^{p\times n}$ being its set of unknown parameters, $F(k)$ denotes the robust adaptive term, and A_d is a user defined diagonal matrix, which must be selected in a way that the eigenvalues of the closed loop system lies within the unit disc [19]. Initial values of the FD estimator are taken to be $\hat{x}(0) = \hat{x}_0$, $\hat{\theta}_d(0) = \hat{\theta}_{d_0}$, $\hat{\theta}_A(0) = \hat{\theta}_{A_0}$, such that $\hat{h}\left(x, u, \hat{\theta}_{d_0}\right) = 0$, $\hat{T}_A(x, u, \hat{\theta}_{A_0}) = [1, ..., 1]^T_{1\times p}$ $\forall x \in X$, $u \in U$.

Remark 2: The proposed observer uses two OLADs to identify additive and multiplicative faults in contrast with other FD schemes which use a single OLAD [7,8,13].

In the proposed FD estimator, NNs are used as the OLADs. Both NN-based OLADs are off prior to the detection of a fault and thus their outputs are zero. Note that as long as the first OLAD is off, all the elements of its output ($\hat{T}_A$) are set to 1. Upon detection of a fault the first OLAD is turned on to estimate the input by assuming that it could be a multiplicative fault. A decision is made to identify the type of fault occurred by monitoring the input residual which is defined as the difference between the actual and estimated input vectors. Depending upon this decision, the appropriate OLAD is left on and the other is switched off. Next the process of detecting a fault is introduced.

Define the detection residual as $e(k) = x(k) - \hat{x}(k)$. Prior to the detection of a fault, the residual dynamics are given by

$$e(k+1) = A_d e(k) + \eta\left(x(k), u(k)\right) \tag{3}$$

which is considered bounded with the appropriate selection of A_d . Now consider a dead-zone operator

$$D[e(k)] = \begin{cases} 0, & \text{if } \|e(k)\| \le \rho \\ e(k), & \text{if } \|e(k)\| > \rho \end{cases}$$

where ρ is the FD threshold. A fault is detected, regardless of its type, when the FD residual exceeds the predefined threshold. However these thresholds must be chosen

carefully in order to minimize false or missed alarms. The selection of the detection threshold is presented next.

Consider the above residual dynamics prior to the occurrence of a fault. If $\hat{x}(0) = x(0)$, the solution to this equation can be obtained as

$$e(k) = \sum_{j=0}^{k-1} A_d^{k-j-1} \left(\eta \left(x(j), u(j) \right) \right) \tag{4}$$

By taking the Frobenius norm, (4) can be represented as

$$\left\| e(k) \right\| = \left\| \sum_{j=0}^{k-1} A_d^{k-j-1} \left(\eta \left(x(j), u(j) \right) \right) \right\|$$

Using Assumption 1, we can find a bound on the residual in the healthy operating conditions, as follows

$$\left\| e(k) \right\| \le \sum_{j=0}^{k-1} \left\| A_d \right\|^{k_d - j - 1} \eta_M \tag{5}$$

Using this bound, a time varying threshold $\rho = \dfrac{\beta \eta_M (1 - \left\| A_d \right\|^k)}{(1 - \left\| A_d \right\|)}$ or a constant threshold $\rho = \dfrac{\beta \eta_M}{(1 - \left\| A_d \right\|)}$ can be determined, where $\beta \ge 1$ is a constant.

When the detection residual exceeds the detection threshold, a fault is declared active through the dead-zone operator and the first OLAD that generates $\hat{T}_A(.)$, is initiated and tuned online using the following update law

$$\hat{\theta}_A(k+1) = \hat{\theta}_A(k) + \alpha_1 \phi_1(k) D\left[e^T(k+1) \right] - \gamma_1 \left\| I - \alpha_1 \phi_1(k) \phi_1^T(k) \right\| \hat{\theta}_A(k) \tag{6}$$

where $\alpha_1 > 0$ is the learning rate, $0 < \gamma_1 < 1$ is the forgetting factor, and $\phi_1(k) = \phi_1\left(x(k), u(k) \right)$ is a basis function such as sigmoid or RBF. Then, the output of the first OLAD that estimates the multiplicative fault function is given by

$$\hat{T}_A(k) = \hat{\theta}_A^T(k) \phi_1\left(x(k), u(k) \right) \tag{7}$$

The input residual is then computed online using actual input and its estimate from the first OLAD. On the other hand, if a fault is identified as additive based on the

input residual, then the first OLAD will be turned off and the second OLAD will be activated and tuned online by using the parameter update law

$$\hat{\theta}_d(k+1)=\hat{\theta}_d(k)+\alpha_2\phi_2(k)e^T(k+1)-\gamma_2\left\|I-\alpha_2\phi_2(k)\phi_2^T(k)\right\|\hat{\theta}_d(k) \tag{8}$$

where $\alpha_2>0$ is the learning rate, $0<\gamma_2<1$ is the forgetting factor, and $\phi_2(k)=\phi_2\left(x(k),u(k)\right)$ is a basis function like sigmoid or RBF. Moreover the output of the second OLAD to estimate the fault function will be given by

$$\hat{h}_d(k)=\hat{\theta}_d^T(k)\phi_2\left(x(k),u(k)\right) \tag{9}$$

A robust adaptive term is also turned on with the second OLAD. The last terms in (6) and (8) overcome the persistent of excitation (PE) condition which is normally utilized to keep the parameters bounded. This assumption is not needed in here and stability is still demonstrated. In the following theorem, conditions for fault detectability are presented.

Theorem 1 (Fault detectability): Consider the nonlinear discrete-time system defined by (1) and the FD estimator described by (2). The fault will be detected, if there exists a time instant k_d, such that the fault function satisfies

$$\frac{1}{2+\left\|A_d^{k_d}\right\|}\left\|\sum_{j=k_0}^{k_0+k_d-1}A_d^{k_0+k_d-j-1}\left(\omega\left(x(j),u_f(j)\right)-\omega\left(x(j),u(j)\right)\right)\right\|>\rho \tag{10}$$

in the case of multiplicative actuator fault or

$$\frac{1}{2+\left\|A_d^{k_d}\right\|}\left\|\sum_{j=k_0}^{k_0+k_d-1}A_d^{k_0+k_d-j-1}\left(\Pi(k-k_0)h\left(x(k),u_f(k)\right)\right)\right\|>\rho \tag{11}$$

in the case of additive fault respectively.

Proof: In the multiplicative fault case, the residual dynamics after the occurrence of fault and prior to the detection of fault is defined by

$$e(k+1)=A_d e(k)+\omega\left(x(k),u_f(k)\right)-\omega\left(x(k),u(k)\right)+\eta\left(x(k),u_f(k)\right) \tag{12}$$

If the fault occurs at time k_0 then after a time instant, k_d, the solution to the above state equation is given by

$$e(k_0+k_d)=A_d^{k_d}e(k_0)+\sum_{j=k_0}^{k_0+k_d-1}A_d^{k_0+k_d-j-1}\left(\omega(x(j),u_f(j))-\omega(x(j),u(j))+\eta(x(j),u(j))\right)$$
$$=A_d^{k_d}e(k_0)+\sum_{j=k_0}^{k_0+k_d-1}A_d^{k_0+k_d-j-1}\left(\eta(x(j),u(j))\right)$$
$$+\sum_{j=k_0}^{k_0+k_d-1}A_d^{k_0+k_d-j-1}\left(\omega(x(j),u_f(j))-\omega(x(j),u(j))\right)$$

By taking the Frobenius norm and applying some mathematical manipulations, we have

$$\left\|e(k_0+k_d)\right\|\geq\left\|\sum_{j=k_0}^{k_0+k_d-1}A_d^{k_0+k_d-j-1}\left(\eta(x(j),u(j))\right)+\sum_{j=k_0}^{k_0+k_d-1}A_d^{k_0+k_d-j-1}\left(\omega(x(j),u_f(j))-\omega(x(j),u(j))\right)\right\|$$
$$-\left\|A_d^{k_d}e(k_0)\right\|$$
$$\geq\left\|\sum_{j=k_0}^{k_0+k_d-1}A_d^{k_0+k_d-j-1}\left(\omega(x(j),u_f(j))-\omega(x(j),u(j))\right)\right\|$$
$$-\left\|\sum_{j=k_0}^{k_0+k_d-1}A_d^{k_0+k_d-j-1}\left(\eta(x(j),u(j))\right)\right\|-\left\|A_d^{k_d}\right\|\left\|e(k_0+k_d)\right\|$$

Now, by using the results we obtained for the detection threshold, we will get

$$\left(1+\left\|A_d^{k_d}\right\|\right)\left\|e(k_0+k_d)\right\|\geq\left\|\sum_{j=k_0}^{k_0+k_d-1}A_d^{k_0+k_d-j-1}\left(\omega(x(j),u_f(j))-\omega(x(j),u(j))\right)\right\|-\rho$$

or equivalently

$$\left\|e(k_0+k_d)\right\|\geq\frac{1}{1+\left\|A_d^{k_d}\right\|}\left(\left\|\sum_{j=k_0}^{k_0+k_d-1}A_d^{k_0+k_d-j-1}\left(\omega(x(j),u_f(j))-\omega(x(j),u(j))\right)\right\|-\rho\right)\quad(13)$$

Thus, the fault will be detected if

$$\frac{1}{1+\left\|A_d^{k_d}\right\|}\left(\left\|\sum_{j=k_0}^{k_0+k_d-1}A_d^{k_0+k_d-j-1}\left(\omega(x(j),u_f(j))-\omega(x(j),u(j))\right)\right\|-\rho\right)>\rho$$

which is equivalent to the detectability condition given in (10) for the case of multiplicative faults.

Now consider the residual dynamics after the occurrence of an additive fault and prior to the detection

$$e(k+1)=A_d e(k)+\Pi(k-k_0)h(x(k),u_f(k))+\eta(x(k),u_f(k))\quad(14)$$

After occurrence of fault and k_d time instants, the solution to the above equation is given by

$$e(k_0+k_d)=A_d^{k_d}e(k_0)+\sum_{j=k_0}^{k_0+k_d-1}A_d^{k_0+k_d-j-1}\left(\eta(x(j),u(j))+\Pi(k-k_0)h(x(k),u_f(k))\right)$$

Similar to the multiplicative fault case, by taking the Frobenius norm and then using the definition of the detection threshold we will get

$$\|e(k_0+k_d)\|\geq\left\|\sum_{j=k_0}^{k_0+k_d-1}A_d^{k_0+k_d-j-1}\left(\Pi(k-k_0)h(x(k),u_f(k))\right)\right\|-\|A_d^{k_d}\|\|e(k_0+k_d)\|-\rho$$

and consequently

$$\|e(k_0+k_d)\|\geq\frac{1}{1+\|A_d^{k_d}\|}\left(\left\|\sum_{j=k_0}^{k_0+k_d-1}A_d^{k_0+k_d-j-1}\left(\Pi(k-k_0)h(x(k),u_f(k))\right)\right\|-\rho\right)\qquad(15)$$

Therefore the detectability condition for the additive fault is defined by (11). The next important issue, after fault detectability condition is the fault detection time. The next theorem gives an upper bound on the detection time.

Theorem 2 (Detection Time): Let the FD estimator in (2) be used to monitor (1). Assume that a fault occurs at $k=k_0$. Also assume there are constants $M_1>0$ and $M_2>0$, such that

$$\omega(x(j),u_f(j))-\omega(x(j),u(j))>M_1\qquad(16)$$

or

$$\Pi(k-k_0)h(x(k),u_f(k))>M_2\qquad(17)$$

for $k_0\leq k\leq k_0+k_{d_{max}}$ for multiplicative or additive faults respectively where $k_{d_{max}}$ is defined by

$$k_{d_{max}}=\max\left(\log_{\|A_d\|}\left(\frac{M_1-2\rho(1-\|A_d\|)}{M_1+\rho(1-\|A_d\|)}\right),\log_{\|A_d\|}\left(\frac{M_2-2\rho(1-\|A_d\|)}{M_2+\rho(1-\|A_d\|)}\right)\right)$$

Then the fault detection time has an upper bound, given by

$$\max(k_{detection})=k_0+k_{d_{max}}$$

Proof: In the case of a multiplicative fault we have

$$\sum_{j=k_0}^{k_0+k_d-1} A_d^{k_0+k_d-j-1}\left(\omega\left(x(j),u_f(j)\right)-\omega\left(x(j),u(j)\right)\right)\geq \sum_{j=k_0}^{k_0+k_d-1}\left\|A_d^{k_0+k_d-j-1}\right\|(M_1)=\frac{1-\left\|A_d^{k_d}\right\|}{1-\|A_d\|}M_1$$

Using the result of theorem 1, a sufficient condition can be found for the fault to be detected, as follows

$$\frac{1-\|A_d\|^{k_d}}{1-\|A_d\|}M_1\geq\rho\left(2+\|A_d\|^{k_d}\right) \tag{18}$$

where k_d is the constant from theorem 1. (18) can be rewritten as

$$\left(M_1+\left(1-\|A_d\|\right)\rho\right)\|A_d\|^{k_d}\leq\left(M_1-2\left(1-\|A_d\|\right)\rho\right)$$

Therefore, an upper bound for k_d in the multiplicative fault case, can be found as

$$k_{d_{max_1}}=\log_{\|A_d\|}\left(\frac{M_1-2\rho(1-\|A_d\|)}{M_1+\rho(1-\|A_d\|)}\right)$$

Similarly, an upper bound for k_d in the additive fault case, is given by

$$k_{d_{max_2}}=\log_{\|A_d\|}\left(\frac{M_2-2\rho(1-\|A_d\|)}{M_2+\rho(1-\|A_d\|)}\right)$$

Since in practice the type of fault is not known, we should take the maximum between $k_{d_{max_1}}$ and $k_{d_{max_2}}$. Then we can find an upper bound for the detection time ($k_{detection}$)

$$\begin{aligned}k_{detection}&\leq k_0+\max(k_{d_{max_1}},k_{d_{max_2}})\\&=k_0+\max\left(\log_{\|A_d\|}\left(\frac{M_1-2\rho\left(1-\|A_d\|\right)}{M_1+\rho\left(1-\|A_d\|\right)}\right),\log_{\|A_d\|}\left(\frac{M_2-2\rho\left(1-\|A_d\|\right)}{M_2+\rho\left(1-\|A_d\|\right)}\right)\right)\end{aligned}$$

Next the performance of the proposed fault diagnosis observer is evaluated on multiplicative faults first and then additive faults.

3.1. MULTIPLICATIVE ACTUATOR FAULT CASE

As mentioned previously, when a fault is detected, only the first OLAD is turned on, while the output of the second OLAD, $\hat{h}_d(.)$, will remain at zero. Therefore after detection, the FD estimator dynamics would be described by

$$\begin{aligned}\hat{x}(k+1) &= A_d\hat{x}(k)+\omega\left(x(k),\hat{u}_f\left(u(k),\hat{T}_A(k)\right)\right)-A_d x(k)\\ \hat{u}_f\left(u(k),\hat{T}_A(k)\right) &= \hat{T}_A\left(x(k),u(k);\hat{\theta}_A(k)\right)u(k)\end{aligned} \tag{19}$$

Consequently, the detection residual dynamics are given by

$$e(k+1) = A_d e(k)+\eta\left(x(k),u_f(k)\right)+\omega\left(x(k),u_f(k)\right)-\omega\left(x(k),\hat{u}_f(k)\right) \tag{20}$$

Define the input residual $\tilde{u}=u-\hat{u}_f$. The next theorem will assure the boundedness of the detection residual dynamics upon detecting a multiplicative fault. Hence the multiplicative fault can be estimated by the first OLAD, which will result in a noticeable difference between the actual and estimated input in a finite time or when the input residual exceeds a user defined threshold. Using this input residual, the fault diagnosis is carried out to identify the fault type.

Theorem 3 (Fault Diagnosis Observer Performance with Multiplicative Actuator Faults): Let the proposed observer defined in (19) be used to monitor the system described by (2), with the first OLAD being turned on upon the detection of a fault. Let the update law in (6) be used to update the unknown parameter vector $\hat{\theta}_A$. In the case of multiplicative faults, the FD residual, $e(k)$, and the parameter estimation errors, $\tilde{\theta}_A(k)$, will be uniformly ultimately bounded (UUB). Moreover, the input residual will exceed the user-defined threshold.

Proof: Consider the following Lyapunov function candidate

$$V = e^T(k)e(k)+tr\left\{\tilde{\theta}_A^T(k)\tilde{\theta}_A(k)\right\}$$

where $\tilde{\theta}_A^T(k)=\theta_A-\hat{\theta}_A(k)$. The first difference of the Lyapunov function is given by

$$\Delta V = \underbrace{\left(e^T(k+1)e(k+1)-e^T(k)e(k)\right)}_{\Delta V_1}+\underbrace{tr\left\{\tilde{\theta}_A^T(k+1)\tilde{\theta}_A(k+1)-\tilde{\theta}_A^T(k)\tilde{\theta}_A(k)\right\}}_{\Delta V_2} \tag{21}$$

By substituting $e(k+1)$ from the error dynamics in ΔV_1, it can be rewritten as

$$\Delta V_1 = \left(A_d e(k)+\tilde{\omega}(k)+\eta\left(x(k),u_f(k)\right)\right)^T\left(A_d e(k)+\tilde{\omega}(k)+\eta\left(x(k),u_f(k)\right)\right)-e^T(k)e(k)$$

where $\tilde{\omega}(k)=\omega\left(x(k),u_f(k)\right)-\omega\left(x(k),\hat{u}_f(k)\right)$.

By using the Cauchy-Schwarz inequality ($(s_1+s_2+\ldots+s_n)^T(s_1+s_2+\ldots+s_n) \le n(s_1^T s_1+s_2^T s_2+\ldots+s_n^T s_n)$) we get

$$\Delta V_1 \le 3e^T(k)A_d^T A_d e(k)+3\tilde{\omega}^T(k)\tilde{\omega}(k)+3\eta^T\left(x(k),u_f(k)\right)\eta\left(x(k),u_f(k)\right) - e^T(k)e(k) \tag{22}$$

Now we substitute $\hat{\theta}_A(k+1)$ from the update law, in ΔV_2

$$\Delta V_2 = tr\left\{\left[\left(1-\gamma_1\left\|I-\alpha_1\phi_1(k)\phi_1^T(k)\right\|\right)\tilde{\theta}_A(k)+\gamma_1\left\|I-\alpha_1\phi_1(k)\phi_1^T(k)\right\|\theta_A-\alpha_1\phi_1(k)e^T(k+1)\right]^T \cdot\left[\left(1-\gamma_1\left\|I-\alpha_1\phi_1(k)\phi_1^T(k)\right\|\right)\tilde{\theta}_A(k)+\gamma_1\left\|I-\alpha_1\phi_1(k)\phi_1^T(k)\right\|\theta_A-\alpha_1\phi_1(k)e^T(k+1)\right] -\tilde{\theta}_A^T(k)\tilde{\theta}_A(k)\right\}$$

Then substituting $e(k+1)$ in the above equation to get

$$\begin{aligned}\Delta V_2 = tr\Big\{&\Big[\left(1-\gamma_1\left\|I-\alpha_1\phi_1(k)\phi_1^T(k)\right\|\right)\tilde{\theta}_A(k)+\gamma_1\left\|I-\alpha_1\phi_1(k)\phi_1^T(k)\right\|\theta_A \\ &-\alpha_1\phi_1(k)\left(A_d e(k)+\tilde{\omega}(k)+\eta\left(x(k),u_f(k)\right)\right)^T\Big]^T \\ &\cdot\Big[\left(1-\gamma_1\left\|I-\alpha_1\phi_1(k)\phi_1^T(k)\right\|\right)\tilde{\theta}_A(k)+\gamma_1\left\|I-\alpha_1\phi_1(k)\phi_1^T(k)\right\|\theta_A \\ &-\alpha_1\phi_1(k)\left(A_d e(k)+\tilde{\omega}(k)+\eta\left(x(k),u_f(k)\right)\right)^T\Big] \\ &-\tilde{\theta}_A^T(k)\tilde{\theta}_A(k)\Big\}\end{aligned}$$

Applying the Cauchy-Schwarz inequality on the above equation yields

$$\begin{aligned}\Delta V_2 \le 5tr\Big\{&\left(1-\gamma_1\left\|I-\alpha_1\phi_1(k)\phi_1^T(k)\right\|\right)^2\tilde{\theta}_A^T(k)\tilde{\theta}_A(k)+\gamma_1^2\left\|I-\alpha_1\phi_1(k)\phi_1^T(k)\right\|^2\theta_A^T\theta_A \\ &-\frac{1}{5}\tilde{\theta}_A^T(k)\tilde{\theta}_A(k)+\alpha_1^2\eta\left(x(k),u_f(k)\right)\phi_1^T(k)\phi_1(k)\eta^T\left(x(k),u_f(k)\right) \\ &+\alpha_1^2 A_d e(k)\phi_1^T(k)\phi_1(k)e^T(k)A_d^T+\alpha_1^2\tilde{\omega}(k)\phi_1^T(k)\phi_1(k)\tilde{\omega}^T(k)\Big\}\end{aligned}$$

Therefore,

$$\begin{aligned}\Delta V_2 \le tr\Big\{&4\tilde{\theta}_A^T(k)\tilde{\theta}_A(k)-10\gamma_1\left\|I-\alpha_1\phi_1(k)\phi_1^T(k)\right\|\tilde{\theta}_A^T(k)\tilde{\theta}_A(k) \\ &+5\gamma_1^2\left\|I-\alpha_1\phi_1(k)\phi_1^T(k)\right\|^2\tilde{\theta}_A^T(k)\tilde{\theta}_A(k)+5\gamma_1^2\left\|I-\alpha_1\phi_1(k)\phi_1^T(k)\right\|^2\theta_A^T\theta_A\Big\} \\ &+5\alpha_1^2 e^T(k)A_d^T A_d e(k)\phi_1^T(k)\phi_1(k)+5\alpha_1^2\tilde{\omega}^T(k)\tilde{\omega}(k)\phi_1^T(k)\phi_1(k) \\ &+5\alpha_1^2\eta^T\left(x(k),u_f(k)\right)\eta\left(x(k),u_f(k)\right)\phi_1^T(k)\phi_1(k)\end{aligned} \tag{23}$$

By using equations (21),(22), and (23), the first difference of the Lyapunov function candidate, ΔV, can be found as

$$\begin{aligned}\Delta V &= \Delta V_1 + \Delta V_2 \\ &\le 3e^T(k)A_d^T A_d e(k) + 3\tilde{\omega}^T(k)\tilde{\omega}(k) + 3\eta^T\left(x(k),u_f(k)\right)\eta\left(x(k),u_f(k)\right) - e^T(k)e(k) \\ &\quad + tr\left\{4\tilde{\theta}_A^T(k)\tilde{\theta}_A(k) - 10\gamma_1\left\|I-\alpha_1\phi_1(k)\phi_1^T(k)\right\|\tilde{\theta}_A^T(k)\tilde{\theta}_A(k)\right. \\ &\qquad \left. + 5\gamma_1^2\left\|I-\alpha_1\phi_1(k)\phi_1^T(k)\right\|^2\tilde{\theta}_A^T(k)\tilde{\theta}_A(k) + 5\gamma_1^2\left\|I-\alpha_1\phi_1(k)\phi_1^T(k)\right\|^2\theta_A^T\theta_A\right\} \\ &\quad + 5\alpha_1^2 e^T(k)A_d^T A_d e(k)\phi_1^T(k)\phi_1(k) + 5\alpha_1^2\tilde{\omega}^T(k)\tilde{\omega}(k)\phi_1^T(k)\phi_1(k) \\ &\quad + 5\alpha_1^2\eta^T\left(x(k),u_f(k)\right)\eta\left(x(k),u_f(k)\right)\phi_1^T(k)\phi_1(k)\end{aligned}$$

Taking the Frobenius norm, the above inequality can be rewritten as

$$\begin{aligned}\Delta V &\le 3A_{d_{max}}^2\left\|e(k)\right\|^2 + 3\left\|\tilde{\omega}(k)\right\|^2 + 5\alpha_1^2\phi_{1_{max}}^2\left\|\tilde{\omega}(k)\right\|^2 - \left\|e(k)\right\|^2 + 5\gamma_1^2\left\|I-\alpha_1\phi_1(k)\phi_1^T(k)\right\|^2\left\|\theta_A\right\|^2 \\ &\quad + 5\alpha_1^2\phi_{1_{max}}^2 A_{d_{max}}^2\left\|e(k)\right\|^2 + 3\left\|\eta\left(x(k),u_f(k)\right)\right\|^2 + 5\alpha_1^2\phi_{1_{max}}^2\left\|\eta\left(x(k),u_f(k)\right)\right\|^2 \\ &\quad + \left(4 - 10\gamma_1\left\|I-\alpha_1\phi_1(k)\phi_1^T(k)\right\| + 5\gamma_1^2\left\|I-\alpha_1\phi_1(k)\phi_1^T(k)\right\|^2\right)\left\|\tilde{\theta}_A(k)\right\|^2 \qquad (24)\end{aligned}$$

Assumption 3 yields $\left\|\tilde{\omega}(k)\right\| \le c_g\left\|u_f(k) - \hat{u}_f(k)\right\|$. Therefore by using Assumptions 4 and 5 to get

$$\begin{aligned}\left\|\tilde{\omega}(k)\right\|^2 &\le c_g^2\left\|diag\left\{T_A\left(x(k),u(k)\right) - \hat{T}_A\left(x(k),u(k);\hat{\theta}_A(k)\right)\right\}u(k)\right\|^2 \\ &\le c_g^2\left\|u(k)\right\|^2\left\|diag\left\{\tilde{\theta}_A(k)\phi_1(k) + \varepsilon_A(k)\right\}\right\|^2 \le 2c_g^2 u_{\max}^2\left(\left\|\tilde{\theta}_A(k)\right\|^2\phi_{1_{\max}}^2 + \varepsilon_{A_M}^2\right)\end{aligned}$$

where $\varepsilon_A(k)$ represents the first OLAD approximation error which is bounded above, i.e. $\left\|\varepsilon_A(k)\right\| \le \varepsilon_{A_M}$. By using the above inequality and the result of Assumption 1 and combining the similar terms in (24), the following inequality is obtained

$$\begin{aligned}\Delta V &\le -\left(1 - (3 + 5\alpha_1^2\phi_{1_{max}}^2)A_{d_{max}}^2\right)\left\|e(k)\right\|^2 \\ &\quad - \left(10\gamma_1\left\|I-\alpha_1\phi_1(k)\phi_1^T(k)\right\| - 5\gamma_1^2\left\|I-\alpha_1\phi_1(k)\phi_1^T(k)\right\|^2 - 4 - 2c_g^2 u_{\max}^2\phi_{1_{\max}}^2\right)\left\|\tilde{\theta}_A(k)\right\|^2 \\ &\quad + (3 + 5\alpha_1^2\phi_{1_{max}}^2)\eta_M^2 + 5\gamma_1^2\left\|I-\alpha_1\phi_1(k)\phi_1^T(k)\right\|^2\left\|\theta_A\right\|^2 + 2c_g^2 u_{\max}^2\varepsilon_{A_M}^2\end{aligned}$$

Hence, the detection residual and parameter estimation errors are uniformly ultimately bounded, if the design parameters are selected by using

$$(3 + 5\alpha_1^2\phi_{1_{max}}^2)A_{d_{max}}^2 < 1 \qquad (25)$$

$$\left\|I-\alpha_1\phi_1(k)\phi_1^T(k)\right\|\left(2-\gamma_1\left\|I-\alpha_1\phi_1(k)\phi_1^T(k)\right\|\right)>\frac{4+2c_g^2u_{\max}^2\phi_{1_{\max}}^2}{5\gamma_1} \tag{26}$$

Moreover, the bounds on the residual and parameter estimation error are given by

$$\left\|e(k)\right\|>\sqrt{\frac{\alpha}{\beta_1}} \quad or \tag{27}$$

$$\left\|\tilde{\theta}_A(k)\right\|>\sqrt{\frac{\alpha}{\beta_2}} \tag{28}$$

where α, β_1, and β_2 are defined by

$$\alpha=(3+5\alpha_1^2\phi_{1_{max}}^2)\eta_M^2+5\gamma_1^2\left\|I-\alpha_1\phi_1(k)\phi_1^T(k)\right\|^2\left\|\theta_A\right\|^2+2c_g^2u_{\max}^2\varepsilon_{A_M}^2,$$

$$\beta_1=1-(3+5\alpha_1^2\phi_{1_{max}}^2)A_{d_{max}}^2,$$

$$\beta_2=10\gamma_1\left\|I-\alpha_1\phi_1(k)\phi_1^T(k)\right\|-5\gamma_1^2\left\|I-\alpha_1\phi_1(k)\phi_1^T(k)\right\|^2-4-2c_g^2u_{\max}^2\phi_{1_{\max}}^2$$

It is obvious that when a multiplicative actuator fault occurs, the difference between u_f and u will not be zero anymore, but will satisfy $u-u_f=(I-T_A)u$, and therefore, $\left\|u-u_f\right\|>T\left\|u\right\|$ where $0<T<\left\|I-T_A\right\|$. Since $\hat{u}_f$ is the estimated value of u_f, the input residual, $\tilde{u}=u-\hat{u}_f$, can be used to determine whether or not a multiplicative actuator fault is present. Upon detection, if the input residual exceeds a predefined threshold, δ, in a finite time, T_δ, then the fault type is identified as multiplicative and the first OLAD is kept online while the second OLAD will never be turned on. By contrast, if the input residual stays below δ, within the interval of T_δ, then the fault type is declared as additive and the second OLAD is turned on and the first one is turned off. If an upper bound for the input u is available, i.e. $\left\|u\right\|\le u_M$, then the threshold δ can be determined by $\delta=\tau_0u_M$ where τ_0 is a positive constant, otherwise a time-varying threshold $\delta=\tau_0\left\|u\right\|$ can be selected.

Remark 3: The time threshold, T_δ, which depends on the time constants of the system and the rate of possible multiplicative faults, is only used to limit the isolation time for the additive faults. This time threshold does not play a significant role in

identifying fault types, since the multiplicative faults can be identified even without using it.

3.2. ADDITIVE FAULT CASE

Since the first OLAD is designed to estimate the multiplicative fault function, it will not be compensating an additive fault. Therefore, in case of an additive fault the estimated input, $\hat{u}_f$, will be close to the actual input, u or the input residual will be below the threshold. So in this case, the second part of Theorem 3 will help identify the fault type after a finite time T_δ, once a fault is detected.

Since in this case only the second OLAD is online. The FD estimator dynamics are described by

$$\hat{x}(k+1) = A_d \hat{x}(k) + \omega\left(x(k), u(k)\right) + \hat{h}_d\left(x(k), u(k); \hat{\theta}_d(k)\right) - A_d x(k) + F(k) \quad (29)$$

where the robust adaptive term, $F(k)$, defined by

$$F(k) = \frac{\hat{\theta}_d^T(k) B}{B^T \hat{\theta}_d(k) \hat{\theta}_d^T(k) B + c} \quad (30)$$

is utilized with the OLAD. Here B is a constant vector and $c > 0$ denotes a positive constant. The following theorem guarantees the performance of the observer with additive faults.

Theorem 4 (Fault Diagnosis Observer Performance with Additive Faults) [13]: Let the proposed observer in (29) be used to monitor the system in (1), with the second OLAD and the robust adaptive term are turned on upon identifying an additive fault. Let the update law in (8) be used to update the unknown parameter set $\hat{\theta}_d$. Then the FD residual, $e(k)$, and the parameter estimation errors, $\tilde{\theta}_d^T(k)$, converge to zero asymptotically.

So far, the detection of a fault and the fault type identification is done. The next section discusses the TTF scheme.

4. PREDICTION SCHEME

Time to failure (TTF) determination is necessary for prognostics. This is also referred to as remaining useful life of the system. Upon detection of a fault, by comparing the estimated parameters obtained from the OLAD to the user defined limits, TTF can be determined [13]. The TTF is defined as the time elapsed when the first parameter reaches its limit. The following theorem provides an analytical formula for finding TTF.

Theorem 5 (TTF Determination): In the presence of multiplicative faults, TTF for the j^{th} parameter of the i^{th} fault, at the k^{th} time instant can be determined using

$$TTF_{i,j}(k) = \frac{\left| \log\left(\left| \frac{\gamma_1 \left\| I - \alpha_1 \phi_1 \phi_1^T \right\| \theta_{A_{i,j\max}} - \alpha_1 \phi_{1_i} e_j^T}{\gamma_1 \left\| I - \alpha_1 \phi_1 \phi_1^T \right\| \hat{\theta}_{A_{i,j}}(k) - \alpha_1 \phi_{1_i} e_j^T} \right| \right) \right|}{\left| \log\left(1 - \gamma_1 \left\| I - \alpha_1 \phi_1 \phi_1^T \right\|\right) \right|} \tag{31}$$

where $\theta_{A_{i,j\max}}$ is the failure limit in terms of maximum value of the system parameter , $\theta_{A_{i,j}}$, and $\hat{\theta}_{A_{i,j}}(k)$ is the estimated system parameter at the time instant k.

Similarly in the presence of additive faults, TTF for the j^{th} parameter of the i^{th} fault, at the k^{th} time instant can be determined using

$$TTF_{i,j}(k) = \frac{\left| \log\left(\left| \frac{\gamma_2 \left\| I - \alpha_2 \phi_2 \phi_2^T \right\| \theta_{d_{i,j\max}} - \alpha_2 \phi_{2_i} e_j^T}{\gamma_2 \left\| I - \alpha_2 \phi_2 \phi_2^T \right\| \hat{\theta}_{d_{i,j}}(k) - \alpha_2 \phi_{2_i} e_j^T} \right| \right) \right|}{\left| \log\left(1 - \gamma_2 \left\| I - \alpha_2 \phi_2 \phi_2^T \right\|\right) \right|} \tag{32}$$

where $\theta_{d_{i,j\max}}$ is the failure limit in terms of maximum value of the system parameter , $\theta_{d_{i,j}}$, and $\hat{\theta}_{d_{i,j}}(k)$ is the estimated system parameter at the time instant *k*.

Proof: Suppose that a fault is detected and identified as multiplicative. Let $a(k) = 1 - \gamma_1 \left\| I - \alpha_1 \phi_1(k) \phi_1^T(k) \right\|$, and let $v(k) = \phi_{1_i}(k) e_j^T(k+1)$. Then the parameter update law in (6) can be rewritten as

$$\hat{\theta}_{A_{i,j}}(k+1) = a(k) \hat{\theta}_{A_{i,j}}(k) + \alpha_1 v(k) \tag{33}$$

which is in the form of the state equation of linear time-varying system with $\hat{\theta}_{A_{i,j}}$ being the state and $v(k)$ being the input.

We know that $0 < a < 1$ and $v(k)$ is bounded since the activation functions are bounded and the boundedness of the residual has been proven earlier. So we can assume that $a(k)$ and $v(k)$ are time invariant. Hence (33) can be rewritten as

$$\hat{\theta}_{A_{i,j}}(k+1) = a\hat{\theta}_{A_{i,j}}(k) + \alpha_1 v \tag{34}$$

which is in form of a linear time-invariant state equation.

At the time of failure, $k_{f_{i,j}}$, estimated parameter will be equal to $\theta_{A_{i,j\max}}$, which means $\hat{\theta}_{A_{i,j}}(k_{f_{i,j}}) = \theta_{A_{i,j\max}}$. Therefore by finding the solution to the linear time-invariant equation in hand, at the time instant $k_{f_{i,j}}$, we get

$$\theta_{A_{i,j\max}} = \hat{\theta}_{A_{i,j}}\left(k_{f_{i,j}}\right) = a^{k_{f_{i,j}}-k}\hat{\theta}_{A_{i,j}}(k) + \alpha_1 \sum_{h=k+1}^{k_{f_{i,j}}} a^{k_{f_{i,j}}-h} v = a^{k_{f_{i,j}}-k}\hat{\theta}_{A_{i,j}}(k) + \alpha_1 v \frac{1-a^{k_{f_{i,j}}-k}}{1-a}$$

Since $TTF_{i,j} = k_{f_{i,j}} - k$, by simple mathematical manipulations, we will have

$$TTF_{i,j} = \frac{\left|\log\left(\left|\frac{(1-a)\theta_{A_{i,j\max}} - \alpha_1 v}{(1-a)\hat{\theta}_{A_{i,j}}(k) - \alpha_1 v}\right|\right)\right|}{\left|\log(a)\right|}$$

By replacing a and v by $(1-\gamma_1\left\|I - \alpha_1\phi_1\phi_1^T\right\|)$ and $\phi_{1_i} e_j^T$ respectively, the desired result will be obtained. The theorem can be proven for the additive fault case by following the same argument as that of the multiplicative case.

Figure 4.1 illustrates the process of finding the TTF after a fault is detected. At each time instant, after calculating the TTF for all of the system parameters, one should take the minimum of time to failure for all of the parameters, to get the overall TTF for the system. This is because the system will be unsafe even if only one of its parameters reaches its limit.

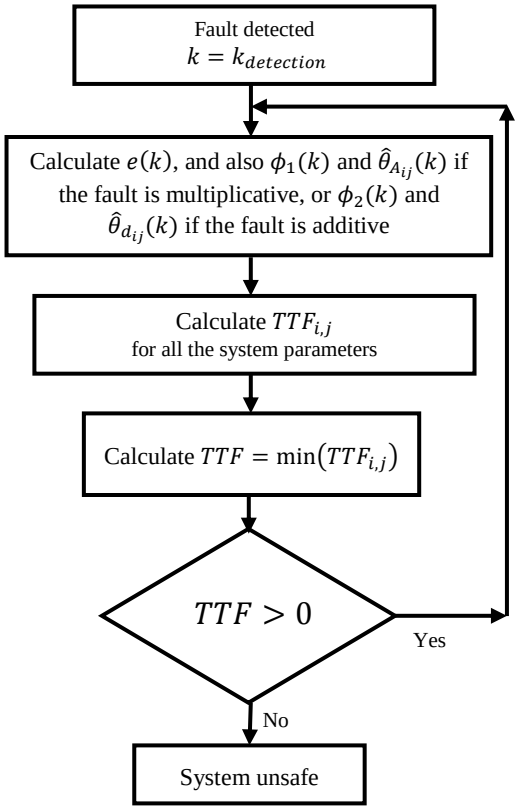

Figure 4.1. Flow chart of the TTF determination

5. SIMULATION RESULTS

In this section, a three-tank water system [21] is used to verify the proposed fault diagnosis and prediction schemes. Figure 5.1 depicts this system consisting of three tanks connected to each other with input pumps on tank 1 and tank 2 respectively and one water outlet on tank 2.

The three-tank system dynamics are described by

$$x(k+1) = \omega(x(k), u(k)) + \eta(x(k))$$

where $x(k) = [x_1(k), x_2(k), x_3(k)]^T$ is the state vector and $\omega(x(k), u(k))$ is the known nonlinear dynamics of the system [21] given by

$$\omega\left(x(k),u(k)\right)=\begin{bmatrix} \frac{T}{A}\left\{-c_1S_p sign\left(x_1(k)-x_3(k)\right)\sqrt{2g\left|x_1(k)-x_3(k)\right|}\right. \\ \left.+u_1(k)\right\}+x_1(k) \\ \\ \frac{T}{A}\left\{-c_3S_p sign\left(x_2(k)-x_3(k)\right)\sqrt{2g\left|x_2(k)-x_3(k)\right|}\right. \\ \left.-c_2S_p\sqrt{2gx_2(k)}+u_2(k)\right\}+x_2(k) \\ \\ \frac{T}{A}\left\{-c_1S_p sign\left(x_1(k)-x_3(k)\right)\sqrt{2g\left|x_1(k)-x_3(k)\right|}\right. \\ \left.-c_3S_p sign\left(x_3(k)-x_2(k)\right)\sqrt{2g\left|x_3(k)-x_2(k)\right|}\right\}+x_3(k) \end{bmatrix} \tag{35}$$

where T is the sampling time chosen to be 0.01 seconds, $A=0.0154m^2$ is the cross section of the tanks, $S_p=5\times10^{-5}\,m^2$ is the cross section of the connecting pipes, $c_1=1, c_2=0.8,$ and $c_3=1$ are the outflow coefficients, and $g=9.8m/s^2$ is the standard gravity. Moreover, $\eta\left(x(k)\right)=\left[10^{-3}\sin(0.7kT)10^{-2}\cos(0.8kT)10^{-1.65}\cos(0.5kT)\right]^T$ represents the modeling uncertainty.

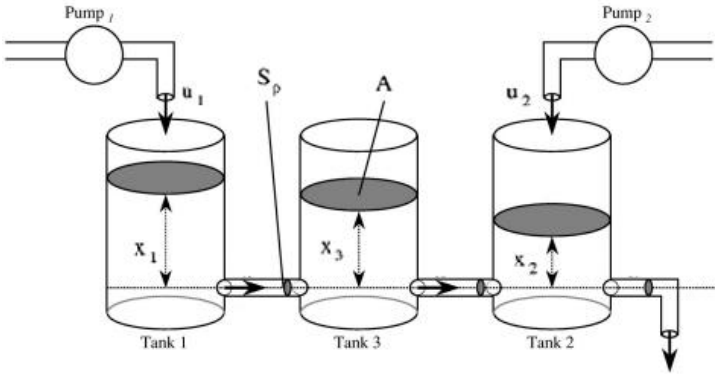

Figure 5.1. Schematic view of the three-tank system

This system is subjected to additive faults which are given in terms of leakage in tank 1 and tank 2 and multiplicative actuator faults which can occur in pump 1 and pump 2. Nevertheless it is assumed that both types cannot be present at the same time. Hence, in this simulation either an additive or a multiplicative fault can occur at time $t_0=25$ sec.

The additive and multiplicative fault functions are described by

$$h\left(x\left(k\right)\right)=\begin{bmatrix}0.0154\left(1-e^{-0.5T\left(k-k_0\right)}\right)\sqrt{2gx_1\left(k\right)}\\0.0182\left(1-e^{-0.2T\left(k-k_0\right)}\right)\sqrt{2gx_2\left(k\right)}\\0\end{bmatrix},$$

$$u_f\left(k\right)=\begin{bmatrix}100\left(1-e^{-0.5T\left(k-k_0\right)}\right)u_1\left(k\right)\\50\left(1-e^{-0.5T\left(k-k_0\right)}\right)u_2\left(k\right)\end{bmatrix}$$

The following FD estimator is used to detect the faults

$$\hat{x}\left(k+1\right)=A_d\hat{x}\left(k\right)+\omega\left(x\left(k\right),\hat{u}_f\left(u\left(k\right),\hat{T}_A(k)\right)\right)+\hat{h}_d\left(x\left(k\right),u\left(k\right);\hat{\theta}_d\left(k\right)\right)-A_dx\left(k\right)$$

where $\hat{u}_f\left(u\left(k\right),\hat{T}_A(k)\right)=\hat{T}_A\left(x\left(k\right),u\left(k\right);\hat{\theta}_A\left(k\right)\right)u(k)$, and $A_d=0.001I_{3\times3}$. The first OLAD output is given by $\hat{T}_A\left(k\right)=\hat{\theta}_A^T\left(k\right)\phi_1\left(V_1x\left(k\right)+B_1\right)$, where $\hat{\theta}_A\in\mathbb{R}^{8\times3}$ is the estimated parameter while $\phi_1\in\mathbb{R}^8$ is a vector of sigmoid functions. The second OLAD output is given by $\hat{h}_d\left(k\right)=\hat{\theta}_d^T\left(k\right)\phi_2\left(x\left(k\right),u\left(k\right)\right)$, where $\hat{\theta}_d\in\mathbb{R}^{8\times3}$ is the estimated parameters while $\phi_2\in\mathbb{R}^8$ is a vector of sigmoid functions. Moreover, V_1, B_1, V_2, and B_2 are selected randomly and the update law parameters are $\alpha_1=0.5,\gamma_1=10^{-4}$ and $\alpha_2=0.1,\gamma_2=10^{-4}$ respectively for the first and second OLADs.

The detection threshold, ρ, is selected to be 0.05 while the identification threshold, δ, is chosen to be 0.01 since the input is upper bounded by $\|u\|\le0.01$**.** Now we need to verify the proposed FD scheme in two cases; first when only multiplicative faults occur at time $t_0=25$ sec, and second when only additive faults occur at time $t_0=25$ sec.

Figure 5.2 shows the norm of the detection residual and the FD threshold when a multiplicative fault occurs. It is clearly seen that the residual remains below the detection threshold prior the occurrence of fault. After the fault occurs, the norm of residual starts to increase and it finally exceeds the threshold at a detection time $t=25.44$ sec. At this point the first OLAD is activated and its update law will estimate the fault function.

About 6 seconds after the detection of the fault, the FD residual falls below the threshold due to the OLAD function approximation property**.** This means that the OLAD has successfully estimated the fault function. As observed in Figure 5.3, the norm of input residual $\tilde{u}(k)$ crosses the identification threshold in the interval of $T_\delta = 2$ sec after the detection, indicating that the fault is of type multiplicative.

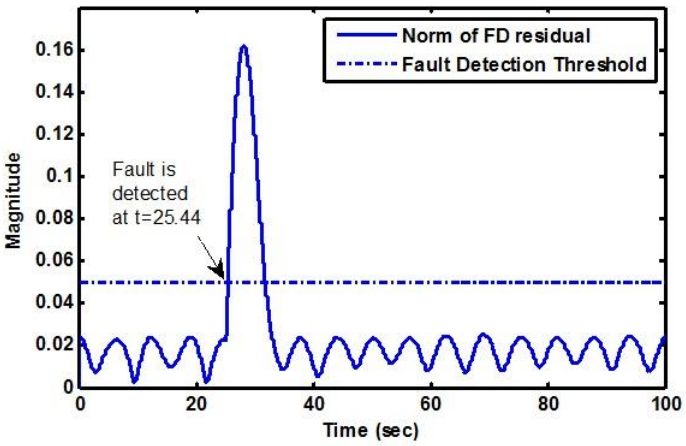

Figure 5.2. Norm of FD residual when the fault is multiplicative

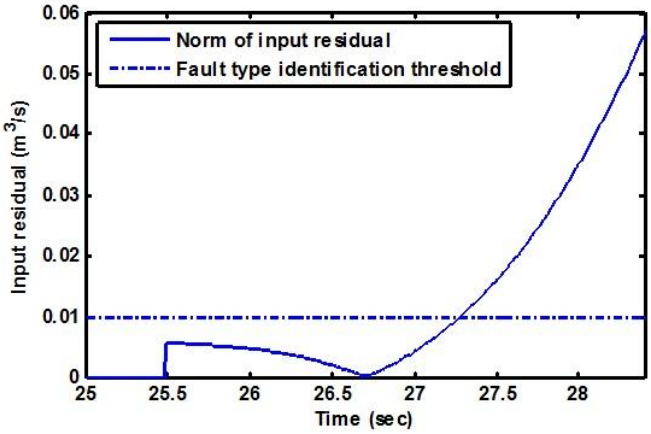

Figure 5.3. Norm of input residual with a multiplicative fault

Figure 5.4 shows the norm of the detection residual along with the detection threshold when additive faults are present. Again the residual remains below the threshold prior to the fault and after that it increases and reaches the detection threshold at time $t = 26.31$ sec. At this point a fault is declared active and the first OLAD is turned

on, but since the fault is additive it cannot estimate the fault function. As seen in Figure 5.5, norm of the input residual $\tilde{u}$ remains below $\delta = 0.01$ within 2 seconds after the detection of fault. Therefore in this case the fault is identified to be additive and the FD estimator uses the second OLAD alone to estimate the fault function upon identifying the fault type. As a consequence, when the second OLAD approximated the fault function, the FD residual converges to zero.

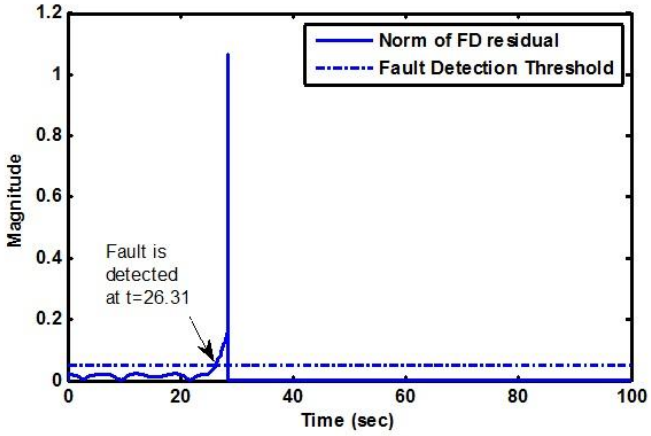

Figure 5.4. FD residual norm with an additive fault

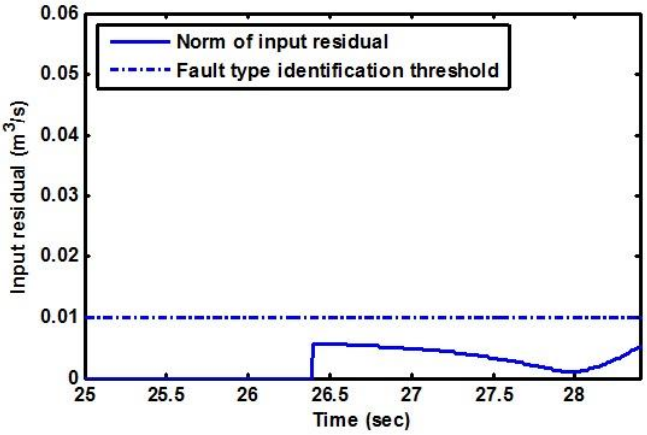

Figure 5.5. Norm of input residual with an additive fault

TTF is determined for each of the multiplicative actuator faults as shown in Figure 5.6 and Figure 5.7. In both cases the initial estimate of TTF is not accurate due to the random selection of weights in the parameter update law. The time of failure is

determined to be at 30.07seconds and 28.30 seconds, for the fault in the first and second inputs respectively. Furthermore Figure 5.8 and Figure 5.9 display TTF estimation versus time for the additive fault in first and second states respectively. The time of failure is determined to be 55.79 seconds for the first fault and 36.63 seconds for the second fault.

This example indicates that the proposed method of fault detection works for both additive and multiplicative fault types, the type of fault can be identified using the proposed method of fault type identification, and furthermore time to failure can be determined using the result of Theorem 5.

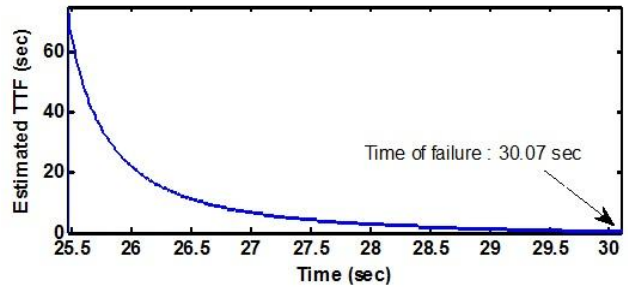

Figure 5.6. TTF determination due to multiplicative fault in input 1

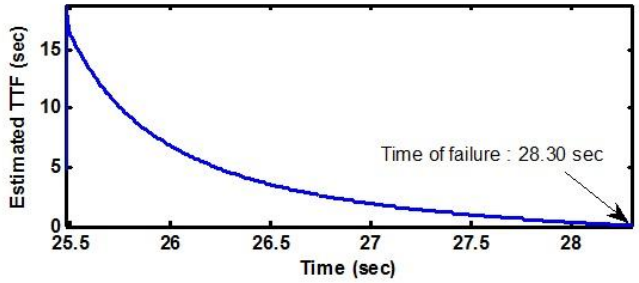

Figure 5.7. TTF determination due to multiplicative fault in input 2

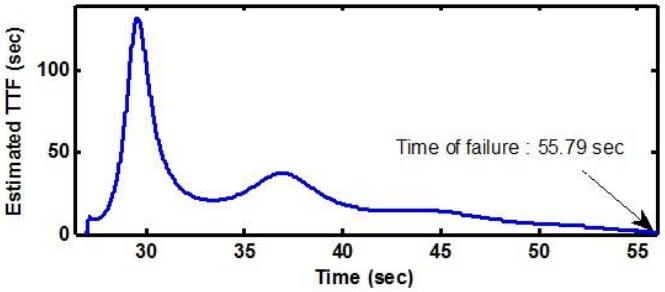

Figure 5.8. TTF determination due to additive fault in state 1

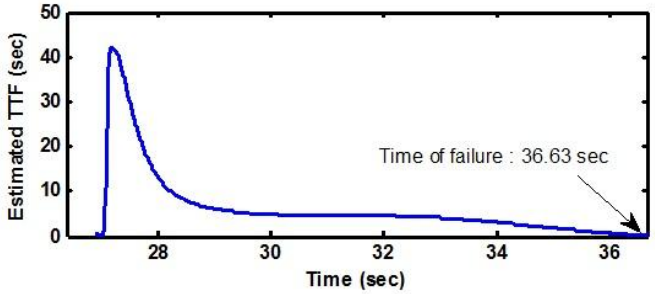

Figure 5.9. TTF determination due to additive fault in state 2

6. CONCLUSIONS

In this paper, a model-based fault detection scheme is proposed that detects both additive system faults and multiplicative actuator fault types, identifies fault type, and performs TTF determination. Input residual generation would help in the detection of multiplicative faults and to identify the type of fault that has occurred in the system. However, this process requires a careful selection of threshold on the input residual. Identification of fault type will help the process of finding the fault location for repair and maintenance purposes. The TTF estimation will in turn improve system availability. The proposed scheme does not need any a priori data or offline training and so it is generic and can be applied to a wide range of systems with a mathematical model available. The only drawback of this scheme is that it requires all the system states to be available. Hence, future work will involve relaxing this assumption.

7. REFERENCES

[1] S. Dash and V. Venkatasubramanian, "Challenges in the industrial applications of fault diagnostic systems," *Comput. Chem. Eng*., vol. 24, no. 2–7, pp. 785–791, 2000.

[2] R. Isermann, "Model-based fault-detection and diagnosis–status and applications," *Annual Reviews in Control*, vol. 29, no.1, pp. 71-85, 2005.

[3] J. Gertler, "Survey of model-based failure detection and isolation in complex plants," *IEEE Control Syst. Mag*., vol. 8, no. 3, pp. 3–11, Dec. 1988.

[4] C. Edwards, S. K. Spurgeon, and R. J. Patton, "Sliding mode observers for fault detection and isolation," *Automatica*, vol. 36, pp. 541-553, 2000.

[5] J. Chen and R. J. Patton, *Robust Model-based Fault Diagnosis for Dynamic Systems*, Kluwer Academic publishers, MA, USA, 1999.

[6] P. M. Frank and L. Keller, "Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy – A survey and some new results," *Automatica*, vol. 26, pp. 459-474, 1990.

[7] M. A. Demetriou and M. M. Polycarpou, "Incipient fault diagnosis of dynamical systems using online approximators," *IEEE Trans. on Automatic Control*, vol. 43, no. 11, pp. 1612-1617, 1998.

[8] M. M. Polycarpou and A. J. Helmicki, "Automated fault detection and accommodation: A learning systems approach," *IEEE Trans. Syst. Man Cybern.*, vol. 25, no. 11, pp. 1447–1458, 1995.

[9] A. Alessandri and T. Parisini, "Model-based fault diagnosis using nonlinear estimators: a neural approach," *Proceedings of the 1997 American Control Conference*, vol. 2, pp. 903-907, 1997.

[10] A. Alessandri, "Fault diagnosis for nonlinear systems using a bank of neural estimators," *Comp. in Industry*, vol. 52, no. 3, pp. 271-289, 2003.

[11] D. Blake and M. Brown, "Simultaneous, multiplicative actuator and sensor fault estimation," *Fuzzy Systems Conference 2007, FUZZ-IEEE*, pp. 1-6, 2007.

[12] C. J. Lopez-Toribio and R. J. Patton, "Fuzzy observers for nonlinear dynamic systems fault diagnosis," *Proc. IEEE Conf. Decision Control*, pp. 84–89, 1998.

[13] B. T. Thumati and S. Jagannathan, "A model based fault detection and prediction scheme for nonlinear multivariable discrete-time systems with asymptotic stability guarantees," *IEEE Transactions on Neural Networks*, vol.21, no. 3, pp. 404-423, 2010.

[14] F. Cacccavale, F. Pierri, and L. Villani, "Adaptive observer for fault diagnosis in nonlinear discrete-time systems," *ASME J. Dyn. Syst. Meas.Control*, vol. 130, no. 2, pp. 1–9, Mar. 2008.

[15] H. P. Huang, C. C. Li, and J. C. Jeng, "Multiple multiplicative fault diagnosis for dynamic processes via parameter similarity measures," *Industrial & Engineering Chemistry Research*, 46, pp. 4517-4530, 2007.

[16] Ch. P. Tan, C. Edwards, "Multiplicative fault reconstruction using sliding mode observers," *Proc. 5th Asian Control Conference*, pp. 957–962, 2004.

[17] H.A. Talebi, and K. Khorasani, "A neural network-based actuator gain fault detection and isolation strategy for nonlinear systems," *IEEE Conference on Decision and Control*, vol. 46, pp. 2614-2619, 2007.

[18] J. Zhang and A. J. Morris, "On-line process fault diagnosis using fuzzy neural networks," *Intelligent Systems Engineering*, pp. 37-47, 1994.

[19] S. Jagannathan, *Neural Network Control of Nonlinear Discrete –time Systems*, CRC publications, NY, 2006.

[20] H. A. Talebi, S. Tafazoli, and K. Khorasani, "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 Trans. Neural Netw*., vol. 20, no. 1,pp. 45–60, 2009.

[21] X. Zhang, M. M. Polycarpou, and T. Parsini, "A robust detection and isolation scheme for abrupt and incipient faults in nonlinear systems," *IEEE Trans. on Automatic Control*, vol. 47, no. 4, pp. 576-593, 2002.

II. An Online Outlier Identification and Removal Scheme for Improving Fault Detection Performance

Hasan Ferdowsi, S. Jagannathan, and M. Zawodniok

Abstract

Measured data or states for a nonlinear dynamic system is usually contaminated by outliers. Identifying and removing outliers will make the data (or system states) more trustworthy and reliable since outliers in the measured data (or states) can cause missed or false alarms during fault diagnosis. In addition, faults can make the system states nonstationary needing a novel analytical model-based fault detection framework. In this paper, an online outlier identification and removal (OIR) scheme is proposed for a nonlinear dynamic system. Since the dynamics of the system can experience unknown changes due to faults, traditional observer-based techniques cannot be used to remove the outliers. The OIR scheme uses a neural network (NN) to estimate the actual system states from measured system states involving outliers. With this method, the outlier detection is performed online at each time instant by finding the difference between estimated and measured states and comparing its median with its standard deviation over a moving time window. The NN weight update law in OIR is designed such that the detected outliers will have no effect on the state estimation which is subsequently utilized for model-based fault diagnosis. In addition, since the OIR estimator cannot distinguish between faulty or healthy operating conditions, a separate model-based observer is designed for fault diagnosis which uses the OIR scheme as a preprocessing unit to improve the fault detection performance. The stability analysis of both OIR and fault diagnosis schemes are introduced. Finally a three-tank benchmarking system and a simple linear system are used to verify the proposed scheme in simulations, and then the scheme is applied on an axial piston pump testbed. The scheme can be applied to nonlinear systems whose dynamics and underlying distribution of states are subjected to change due to both unknown faults and operating conditions.

1. INTRODUCTION

Outliers are present in data sets of practical industrial systems. By definition [1], an outlier is an observation which deviates sufficiently from other observations thus creating suspicion that it was from a different system. In industrial systems, outliers can appear in the measured data. If measured data from a system is contaminated by outliers, processing the data becomes difficult since these outliers can render inaccurate decisions during fault diagnosis. In many cases, the underlying distribution of the measured data can change due to outliers.

On the other hand, due to the high risk of component and system failures, reliable fault diagnosis schemes are required to guarantee safe system operation even in the presence of uncertainties, outliers, and faults. A reliable fault detection scheme is the one that can detect faults at an early stage, without missed or false alarms before the root cause analysis.

Recently the topic of fault detection and diagnosis has attracted a number of researchers around the world. Fault detection is performed data-driven [2], model-based [3] or a combination of both [4, 5]. Several model-based fault detection techniques have been developed in the past decade [6-10]. However, even the best fault detection schemes can become unreliable in the presence of data corrupted with outliers since outliers can cause false alarms.

Several outlier identification and removal schemes have been proposed in the literature such as distribution-based [11], distance-based [12, 13], clustering [14, 15], and density-based methods [16]. Also surveys of different outlier detection methods are given in [17, 18]. However these methods are data-driven and work offline. An online outlier detection scheme is a prerequisite for improving the performance of the model-based fault detection scheme where system states are utilized. Therefore, several online outlier removal methods have also been developed for linear systems with known dynamics by using Kalman filter and its variations [19-21] by assuming that the system states and measurement noise belong to the Gaussian distribution. However, due to changes in the operating conditions and presence of faults, the underlying distribution of states is not necessarily fixed and therefore nonstationary. Therefore a novel scheme to detect the

presence and the removal of outliers from system states is needed for both linear and nonlinear systems in nonstationary environments.

Since the system states are considered to be contaminated with noise and outliers, the main objective of this paper is to develop an online outlier identification and removal (OIR) scheme for system states prior to fault detection stage in contrast with traditional model-based fault detection framework where outlier-free state assumption is made [22]. The robustness and reliability of the model-based fault detection scheme is evaluated by using detection rate, missed and false alarms with and without outliers. In model-based fault detection, data points are system states or outputs.

Since in practice outlier-free system states are not available, they need to be estimated and then compared to the measured states to detect outliers. Here, traditional observers cannot be used because the system dynamics are not known and subjected to unknown faults. Therefore a two-layer feedforward neural network (NN) is utilized to estimate the actual system states and at the same time to identify and remove the outliers. The NN outputs are the estimated states filtered for noise and outliers. At each time instant, the estimated state vector from the OIR scheme is calculated for the next instant of time and compared to the measured states to generate the state estimation error.

Next, median and standard deviation of the state estimation error in a limited time window are found. If the state estimation error and the calculated median are both higher than three standard deviations, an outlier is detected**.** A novel NN weight update law is derived by using the state estimation error. In order to prevent an update of the NN weights in the state estimator in response to an outlier, a variable learning rate is selected such that it takes on a zero value when an outlier is detected. The stability of the state observer utilized for OIR is discussed in the paper.

For the purpose of fault detection, a different observer with known nominal dynamics of the system is introduced. Since the OIR scheme estimates the known system dynamics, uncertainties, and fault function, it cannot be used as a fault detection observer. Moreover by using a second observer for fault detection, the fault function can be approximated for isolation and prognostics. Therefore, an observer-based fault detection scheme is introduced that uses the estimated outlier-free state vector instead of the measured state vector. A fault is detected by comparing the observed states with the

outlier-free system state vector. Upon detection, an online approximator is activated to estimate the fault dynamics. The performance of the fault detection scheme is evaluated with and without the proposed outlier scheme. Again note that the state estimation for the OIR scheme is different than the one used for fault diagnosis.

Therefore the contributions of this paper involve the development of an OIR scheme which can operate online in contrast with data-based methods[11-16], and can be applied to both linear and nonlinear systems in nonstationary environments in contrast with existing Kalman filter-based schemes [19-21]. Since the proposed NN estimator is quite generic and does not use the system representation or model, it is useful even when the system dynamics are not known. In other words, the OIR scheme can be utilized both for data driven and model-based fault diagnosis schemes.

Moreover, a model-based fault detection scheme which uses the estimated outlier-free states instead of the actual measured states of the system is presented and the stability analysis of the proposed fault diagnosis scheme is included when the underlying distribution of states are nonstationary in contrast with all the available model-based fault diagnosis [3-8]**.** This requires a complete novel analytical framework.

To verify the performance and effectiveness of the proposed outlier removal technique and observe its effect on fault detection process, a three-tank water system is used. A fault is seeded in one of the tanks and outlier removal is performed on both healthy and faulty data. It is shown that fault detection can only provide reliable results when the outliers are removed from the measured data. Also a linear example is used to compare the proposed scheme with a Kalman filter-based method. Simulations have been repeated for a significant number of times to evaluate the proposed scheme in different cases of random noise and outliers. Further, an experimental study has been conducted on an axial piston pump testbed in healthy operating conditions and it is shown that the measured outlet pressure involves several outliers which will trigger false alarms during fault detection. The outliers are shown to be removed successfully from the measured outlet pressure by using this online OIR scheme.

This paper is organized as follow: Section II introduces the system description and required assumptions. Section III presents the outlier detection and removal

technique while fault detection scheme is introduced in Section IV. Section V discusses simulation and experimental results.

2. SYSTEM DESCRIPTION

Consider the nonlinear discrete-time system described by the following state space representation

$$x(k+1)=\omega\big(x(k),u(k)\big)+\eta\big(x(k),u(k)\big)$$

where $u \in \mathbb{R}^m$ is the control input vector, $x \in \mathbb{R}^n$ is the system state vector, $\omega: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ represents the known nonlinear system dynamics, and $\eta: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ represents the system uncertainties.

Now consider the nonlinear system with a fault as

$$x(k+1)=\omega\big(x(k),u(k)\big)+\eta\big(x(k),u(k)\big)+\Pi(k-k_0)h\big(x(k),u(k)\big) \qquad (1)$$

where $h(x(k),u(k))$ represents a vector of possible fault dynamics. The time profile of a fault is given by $\Pi(k-k_0)$. The time profile $\Pi(k-k_0)$ is modeled by $\Pi(k-k_0) = diag\{\Omega_1(k-k_0),\Omega_2(k-k_0),\dots,\Omega_n(k-k_0)\}$ where

$$\Omega_i(k-k_0)=\begin{cases}0 & ,\ if\ \tau<0\\ 1-e^{-\bar{\kappa}_i\tau} & ,\ if\ \tau\geq 0\end{cases} \qquad for\ i=1,\dots,n$$

is the time profile variable and $\bar{\kappa}_i$ is an unknown constant that represents the rate at which a fault occurs. A larger value of $\bar{\kappa}_i$ indicates that it is an abrupt fault. The use of such time profiles is common in fault diagnosis literature [6].

Note that this fault will definitely change the system dynamics and might even change the underlying distribution of system states. Normally in the literature [6-9], it is assumed that the states are free with noise and does not change its underlying distribution which is not practical. Consequently, Kalman filters [23] cannot be used to eliminate noise and outliers from the measured states because they require system states and noise to have a fixed distribution and also require the exact system representation which is not available in our case due to unknown fault. In this paper, a NN-based approach will be taken with appropriate selection of NN weights such that this assumption is relaxed.

For the purpose of monitoring the system and performing fault detection, state measurements are required. Usually, the measured system states are contaminated with noise and outliers. The measured state vector $y(k)$ can be represented by

$$y(k) = x(k) + v(k)$$

where $v(k)$ includes measurement noise and outliers which is considered bounded above such that $\|v(k)\| \leq v_M$. The distribution of the measurement noise can change over time. According to the definition [1], if the measured states $y(k)$ deviates significantly from the actual system states $x(k)$, the data point is said to be an outlier. Model-based fault detection schemes cannot distinguish a residual increasing due to a fault or an outlier. Therefore, outliers in the measured states can cause false alarms during fault detection and diagnosis. This fact clearly emphasizes the importance of detection and removal of outliers before fault detection. The following standard assumptions are needed in order to proceed.

Assumption 1: The modeling uncertainty is bounded, i.e. $\|\eta(x(k), u(k))\| \leq \eta_M$, $\forall (x, u) \in (\chi \times U)$, where η_M is a positive known constant.

Remark 1: Assumption 1 is needed to distinguish between faults and system uncertainties and to analytically determine the fault detection threshold.

Assumption 2: The nonlinear system dynamics $\omega(x, u)$ is Lipschitz in x, i.e., $\|\omega(x_1, u) - \omega(x_2, u)\| \leq c_g \|x_1 - x_2\|$, where $c_g > 0$ is the Lipschitz constant.

Remark 2: This assumption is only required for the fault detection part, mainly because the estimated outlier-free states are used in the proposed FD estimator instead of the actual system states. This assumption has been used in other articles on fault diagnosis [6, 24] where the entire state vector is not available and estimated states have to be used in the estimator dynamics instead of actual system states.

Assumption 3: The functions $\omega(x(k), u(k))$, $\eta(x(k), u(k))$, and $h(x(k), u(k))$ can be expressed as nonlinear in the unknown parameters (NLIP), thus can be approximated by two-layer neural networks with bounded weights and approximation errors.

Next the proposed outlier detection scheme is introduced.

3. OUTLIER IDENTIFICATION AND REMOVAL SCHEME

The main objective of this work is to design an outlier scheme which can detect, identify and remove the outliers online, before an outlier triggers a false alarm during fault diagnosis. Therefore, the outlier detection must be performed online and prior to fault detection and root cause analysis.

According to Chebschev's theorem and outlier detection method [25], almost all the observations in a data set of system states will fall into the interval $[\mu - 3\sigma, \mu + 3\sigma]$, where μ and σ are the mean and standard deviation of the data set respectively, and the data points outside this interval are declared outliers. If the distribution of the actual system states was fixed over time, traditional outlier detection methods [25, 26] can be employed whereas for the present scenario, these methods cannot be utilized. Now initially assume that the measured system state vector y has following fixed distribution

$$y(k) \sim \mathcal{N}\left(x(k), \sigma^2\right)$$

where $\mathcal{N}(\mu, \sigma^2)$ is a Gaussian distribution with mean μ and variance σ^2. In this case an outlier can be defined as a point where $|y(k) - x(k)| > 3\sigma$ where $|\cdot|$ is the absolute value operator. If the mean value of the actual states is equal to μ, then this definition can be rephrased as $|y(k) - \mu| > 3\sigma$. But this method is offline and also it cannot be used when the system states do not have a fixed distribution. In this work, we have assumed that the system is subjected to a fault, which can change the nominal dynamics of the system as well as the underlying distribution of the states.

In order to develop a method of online outlier detection for a system with changing dynamics, we will investigate the measured state vector in a fixed time window, assuming that the measurement noise has a fixed distribution over each of these small time windows. Suppose that the state vector at time instant k is being investigated and consider a finite window of time with length p in which the measured state vector is available, i.e. $\{y(k-p+1), \ldots, y(k-1), y(k)\}$.

If the outlier-free state vector in the current window is available, then the difference between actual and measured state vector can be calculated by $\delta = y - x$ and its mean and variance over the selected time interval can be found by

$$\mu(k)=\frac{1}{p}\sum_{j=0}^{p-1}(\delta(k-j)),\ \sigma^2(k)=\frac{1}{p}\sum_{j=0}^{p-1}(\delta(k-j)-\mu(k))^2$$

Since the distribution of measurement noise is normal, assuming that its variance is constant within the considered time window, an outlier can be detected at time k, when $|\delta(k)-\mu(k)|>3\sigma(k)$. Although this is an online outlier detection method which can also handle the changes in the system dynamics, it is impossible to implement since the outlier free system states are not available in practice and the only available data would be the measured states contaminated with noise and outliers.

To overcome this issue, an estimator will be proposed to estimate the unknown system states by assuming that the states are available for measurement. In the literature, the outlier removal is traditionally done without an estimator while an observer is normally utilized for model-based fault detection and not for outlier removal. In contrast, by using an observer, we are estimating the actual states and also performing outlier removal.

If the system dynamics was known, it could be used to construct an observer to estimate the system states, similar to Kalman filter-based outlier detection methods. But this is only possible when the system is working in healthy conditions with known dynamics. In our case, the system is subjected to unknown changes like faults. So an estimator which is able to approximate the system states without using the system dynamics is required. To construct this online approximator and its learning mechanism, initially we consider the case when the measured data does not have outliers, and then the general case will be investigated. Since$\omega\big(x(k),u(k)\big)$, $\eta\big(x(k),u(k)\big)$, $h\big(x(k),u(k)\big)$ are all smooth functions, $x(k+1)$ in equation (1) can be approximated by a two layer NN, if x_m does not involve outliers. So $x(k+1)$ can be written as

$$x(k+1)=W^T(k)\psi\big(x(k),u(k)\big)+\mu(k)$$

where $W\in\mathbb{R}^{q\times n}$ is the unknown parameter matrix which will change when a fault occurs in the system or the model parameters change due to shift in the operating conditions, which can also change the distribution of the states, $\psi\big(x(k),u(k)\big)$ is a basis function like sigmoid, and $\mu(k)$ is the approximation error which is bounded by μ_M [27].

Now let the estimated states be denoted as x_s and consider the NN output as

$$x_s(k+1) = \hat{W}^T(k)\psi\left(x_s(k), u(k)\right) \tag{2}$$

where $\hat{W} \in \mathbb{R}^{q\times n}$ represents the unknown weights of the output layer of NN. Now an update law for training $\hat{W}$ is required. Define the state estimation error $\tilde{x}(k) = x(k) - x_s(k)$ and parameter estimation error$\tilde{W}(k) = W(k) - \hat{W}(k)$. When there is no noise and outliers in the measured states, which means y is equal to x at all times, the weight update law can be selected as

$$\begin{aligned}\hat{W}(k+1) = \hat{W}(k) &+ \beta\psi\left(x_s(k), u(k)\right)e^T(k+1) \\ &- \gamma\left\|I - \beta\psi\left(x_s(k), u(k)\right)\psi^T\left(x_s(k), u(k)\right)\right\|\hat{W}(k)\end{aligned} \tag{3}$$

where $\beta > 0$ is a constant learning rate, $0 < \gamma < 1$, and $e(k+1) = y(k+1) - x_s(k+1)$. Note that the NN weights are updated by the difference between measured states and estimated states, because the actual system states are not available. Then the state estimation error can be written as

$$\tilde{x}(k+1) = W^T(k)\tilde{\psi}(k) + \tilde{W}^T(k)\psi\left(x_s(k), u(k)\right) + \mu(k)$$

where $\tilde{\psi}(k) = \psi\big(x(k), u(k)\big) - \psi\big(x_s(k), u(k)\big)$.

Remark 3: Instead of the measured state vector y, the delayed output of the NN ($x_s(k)$) along with the input vector $u(k)$are used as NN inputs, in order to prevent the outliers in measured data from affecting the state estimates whereas y is only used for updating the NN weights. Later on, the weight update law in (3) will also be modified in order to cancel the effect of outliers on the NN weights.

By choosing the following Lyapunov function candidate

$$V = \tilde{x}^T(k)\tilde{x}(k) + \frac{1}{\beta}tr\{\tilde{W}^T(k)\tilde{W}(k)\}\ ,$$

it can easily be shown that state and parameter estimation errors will be uniformly ultimately bounded. However, measured states involve outliers, so this approach cannot be utilized since the outlier-free state vector is not available. In other words, when y is contaminated with outliers and this measured data is used to update the NN weights in

(3), the actual states will not be estimated correctly while the outlier detection will also be unreliable.

To solve this problem the outlier detection and state estimation processes will be combined to properly detect outliers and design a new weight update law that is not affected by the outliers. The parameter update law in (3) is modified by using a variable learning rate whose value will be zero when an outlier is detected at time $(k+1)$ and not zero otherwise. Suppose that $y(k+1)$ is an outlier. In this case $e(k+1)$ which is used to update the parameters will be large, even if the weights are close to their desired values and $x_s(k+1)$ is close to its desired value $x(k+1)$. To prevent the NN weights to be updated by an outlier at this time instant, the variable leaning rate $\beta(k+1)$ used in the update law should take zero value.

Since $\widehat{W}(k)$ is available at the time instant k, $x_s(k+1)$ can be calculated and used for outlier detection before updating the weights. To perform the outlier detection on $y(k+1)$, again consider a finite window of time with length p. The median value of $\|e\| = \|y - x_s\|$ (where $\|\cdot\|$ is the norm operator) in a window ending at time $(k+1)$ is defined by

$$M(k+1) = Median\left\{\left\|e(k-p+2)\right\|,\ldots,\left\|e(k)\right\|,\left\|e(k+1)\right\|\right\}$$

and the standard deviation in the same time window is defined by

$$\sigma(k+1) = \sqrt{Var\left\{\left\|e(k-p+2)\right\|,\ldots,\left\|e(k)\right\|,\left\|e(k+1)\right\|\right\}}$$

Similar to the first case, it can be assumed that the variance of the measurement noise is constant within the time window. Therefore, a threshold value of three times the standard deviation is used to detect the outliers. Because of the limited time window, the mean value of the data set inside a window might be significantly affected even by a single outlier, which might increase the probability of a false or missed alarm. In contrast, median value is not easily affected by a single outlier. Therefore if mean value is used for outlier detection, the unwanted change in the mean value can definitely degrade the performance of outlier detection process. This simple example clarifies the reason why median is used in this outlier detection scheme instead of mean value.

Thus, median value is used instead of mean value to overcome this problem. Finally the data point at time $(k+1)$ is considered an outlier if

$$\left|\left\|e(k+1)\right\| - M(k+1)\right| > 3\sigma(k+1)$$

Now that the outlier detection is performed for the data set comprised of system states at time $(k+1)$, we need to construct an analytical formula to find the variable learning rate $\beta(k+1)$ based on whether or not an outlier exists at this time instant. The idea is to reduce the learning rate, preferably to zero, when an outlier is detected at time $(k+1)$, in order to prevent the NN weights from getting updated by an outlier. Also, the learning rate needs to be small when outliers are not present with relatively large amplitude noise.

For this purpose, define the function $S(\cdot)$ as

$$S(z) = \begin{cases} \left(1-z^2\right)^2 & for\,|z|<1 \\ 0 & otherwise \end{cases}$$

This bell-shaped function achieves its maximum at $z = 0$ and takes zero value when $z \geq 1$. This function is utilized to construct the robust variable learning rate given by

$$\beta(k+1) = \beta_M S\left(\frac{1}{3\sigma(k+1)}\left|\left\|e(k+1)\right\| - M(k+1)\right|\right)$$

where β_M is the maximum possible learning rate parameter which keeps the estimator stable. Larger noise amplitude will result in larger values for $|\|e(k+1)\| - M(k+1)|$, thus smaller values for the learning rate. Particularly when $|\|e(k+1)\| - M(k+1)| > 3\sigma(k+1)$(which means an outlier is detected at time$(k+1)$) $\beta(k+1)$ will automatically be set to zero, so the weights will not be updated upon detecting an outlier. Further, considering the definition of outliers, it can be inferred that an outlier is detected at time k if $\beta(k) = 0$.

Finally, the proposed parameter update law can be represented by

$$\begin{aligned} \hat{W}(k+1) = \hat{W}(k) &+ \beta(k+1)\psi\left(x_s(k), u(k)\right)e^T(k+1) \\ &- \gamma\left\|I - \beta(k+1)\psi\left(x_s(k), u(k)\right)\psi^T\left(x_s(k), u(k)\right)\right\|\hat{W}(k) \end{aligned} \qquad (4)$$

The definition of the learning rate implies that, if $e(k+1)$ is relatively close to $M(k+1)$, then the corresponding learning rate $\beta(k+1)$ which appears in the parameter update law, will be close to maximum possible learning rate. Whereas, if $e(k+1)$ is largely deviated from $M(k+1)$ then the corresponding learning rate, will be zero or close to

zero. This means that measurement noise cannot make significant change on the NN weights while the effect of outliers on the weight update law is completely eliminated. In the following theorem, the performance of state estimation with the proposed outlier detection and removal scheme is discussed.

Theorem 1: Let an adaptive observer in (2) be used to estimate the state vector of system (1) when the measured state vector y is contaminated with outliers. Then the state estimation error, $\tilde{x}(k) = x(k) - x_s(k)$, and the NN weight estimation error $\tilde{W}(k)$ are uniformly ultimately bounded (UUB) in the mean if the user-defined variables are selected such that

$$3W_{\max}^2 c_\psi^2 \left(1 + 2\beta_M^2 \psi_{\max}^2\right) < 1$$

$$2\gamma \left\| I - \beta(k+1)\hat{\psi}\left(k\right)\hat{\psi}^T\left(k\right)\right\| > 2\beta_M^2 \psi_{\max}^4 + 3\psi_{\max}^2 + 2/3 + \gamma^2 \left\| I - \beta(k+1)\hat{\psi}\left(k\right)\hat{\psi}^T\left(k\right)\right\|^2$$

Proof: Consider the following Lyapunov function

$$V = E\left(tr\left\{\tilde{x}^T\left(k\right) I \tilde{x}\left(k\right)\right\}\right) + \frac{1}{3} E\left(tr\left\{\tilde{W}^T\left(k\right)\tilde{W}(k)\right\}\right)$$

where E(.) is the expectation operator. The first difference of this function is given by

$$\Delta V = E\left(tr\left\{\tilde{x}^T\left(k+1\right) I \tilde{x}\left(k+1\right) - \tilde{x}^T\left(k\right) I \tilde{x}\left(k\right)\right\}\right)$$
$$+ \frac{1}{3} E\left(tr\left\{\tilde{W}^T\left(k+1\right)\tilde{W}\left(k+1\right) - \tilde{W}^T\left(k\right)\tilde{W}\left(k\right)\right\}\right)$$

By substituting $\tilde{x}(k+1)$ from the state estimation error equation and $\tilde{W}(k+1)$ from the update law in the above equation and applying Cauchy-Schwarz inequality, it can be shown that

$$\Delta V \le E\left(tr\left\{3\hat{\psi}^T\left(k\right)\tilde{W}\left(k\right)\tilde{W}^T\left(k\right)\hat{\psi}\left(k\right)\right.\right.$$
$$\left. + 3\tilde{\psi}^T\left(k\right)W(k)W^T(k)\tilde{\psi}\left(k\right) + 3\mu^T(k)\mu(k) - \tilde{x}^T\left(k\right)\tilde{x}\left(k\right)\right\}\Big)$$
$$+ \frac{1}{3} E\Big(tr\Big\{3\left(1 - \gamma\left\| I - \beta(k+1)\hat{\psi}\left(k\right)\hat{\psi}^T\left(k\right)\right\|\right)^2 \tilde{W}^T\left(k\right)\tilde{W}\left(k\right)$$
$$+ 3\beta^2\left(k+1\right)e(k+1)\hat{\psi}^T\left(k\right)\hat{\psi}\left(k\right)e^T(k+1)$$
$$\left. + 3\gamma^2\left\| I - \beta(k+1)\hat{\psi}\left(k\right)\hat{\psi}^T\left(k\right)\right\|^2 W^T\left(k\right)W\left(k\right) - \tilde{W}^T\left(k\right)\tilde{W}\left(k\right)\right\}\Big)$$

$$\le 3\psi_{\max}^2 E\left(\left\|\tilde{W}(k)\right\|^2\right)+3E\left(tr\left\{\tilde{\psi}^T(k)W(k)W^T(k)\tilde{\psi}(k)\right\}\right)+3E\left(\left\|\mu(k)\right\|^2\right)-E\left(\left\|\tilde{x}(k)\right\|^2\right)$$
$$+\frac{1}{3}E\Big(tr\Big\{-6\gamma\left\|I-\beta(k+1)\hat{\psi}(k)\hat{\psi}^T(k)\right\|\tilde{W}^T(k)\tilde{W}(k)$$
$$+3\gamma^2\left\|I-\beta(k+1)\hat{\psi}(k)\hat{\psi}^T(k)\right\|^2\tilde{W}^T(k)\tilde{W}(k)$$
$$+6\beta^2(k+1)\tilde{x}(k+1)\hat{\psi}^T(k)\hat{\psi}(k)\tilde{x}^T(k+1)$$
$$+6\beta^2(k+1)v(k+1)\hat{\psi}^T(k)\hat{\psi}(k)v^T(k+1)$$
$$+3\gamma^2\left\|I-\beta(k+1)\hat{\psi}(k)\hat{\psi}^T(k)\right\|^2W^T(k)W(k)+2\tilde{W}^T(k)\tilde{W}(k)\Big\}\Big)$$

where $\hat{\psi}(k)=\psi\big(x_s(k),u(k)\big)$. Note that $E\big(v^T(k+1)v(k+1)\big)=E(\|v(k+1)\|^2)\le v_M^2$. Assuming that the basis function $\psi(.)$ is a Lipschitz function with the Lipschitz constant c_ψ, leads to

$$\Delta V\le 3\psi_{\max}^2 E\left(\left\|\tilde{W}(k)\right\|^2\right)+3W_{\max}^2c_\psi^2E\left(\left\|\tilde{x}(k)\right\|^2\right)+3\mu_M^2-E\left(\left\|\tilde{x}(k)\right\|^2\right)+\frac{2}{3}E\left(\left\|\tilde{W}(k)\right\|^2\right)$$
$$-2\gamma\left\|I-\beta(k+1)\hat{\psi}(k)\hat{\psi}^T(k)\right\|E\left(\left\|\tilde{W}(k)\right\|^2\right)$$
$$+\gamma^2\left\|I-\beta(k+1)\hat{\psi}(k)\hat{\psi}^T(k)\right\|^2E\left(\left\|\tilde{W}(k)\right\|^2\right)$$
$$+\gamma^2\left\|I-\beta(k+1)\hat{\psi}(k)\hat{\psi}^T(k)\right\|^2W_{\max}^2+2\beta_M^2\psi_{\max}^2v_M^2$$
$$+2\beta_M^2\psi_{\max}^2\left(\psi_{\max}^2E\left(\left\|\tilde{W}(k)\right\|^2\right)+3W_{\max}^2c_\psi^2E\left(\left\|\tilde{x}(k)\right\|^2\right)+3\mu_M^2\right)$$

Where $W_{max}$and $\psi_{max}$are the maximum norm values of W and ψ. Substituting $\tilde{x}(k+1)$ from the state estimation error equation and combining similar terms yields

$$\Delta V\le-\left(1-3W_{\max}^2c_\psi^2\left(1+2\beta_M^2\psi_{\max}^2\right)\right)E\left(\left\|\tilde{x}(k)\right\|^2\right)$$
$$-\Big(2\gamma\left\|I-\beta(k+1)\hat{\psi}(k)\hat{\psi}^T(k)\right\|-2\beta_M^2\psi_{\max}^4-3\psi_{\max}^2$$
$$-\frac{2}{3}-\gamma^2\left\|I-\beta(k+1)\hat{\psi}(k)\hat{\psi}^T(k)\right\|^2\Big)E\left(\left\|\tilde{W}(k)\right\|^2\right)$$
$$+\gamma^2\left\|I-\beta(k+1)\hat{\psi}(k)\hat{\psi}^T(k)\right\|^2W_{\max}^2+2\beta_M^2\psi_{\max}^2v_M^2+3\mu_M^2\left(1+2\beta_M^2\psi_{\max}^2\right)$$

Define b_1, b_2, and D as

$$B_1=1-3W_{\max}^2c_\psi^2\left(1+2\beta_M^2\psi_{\max}^2\right)$$
$$B_2=2\gamma\left\|I-\beta(k+1)\hat{\psi}(k)\hat{\psi}^T(k)\right\|-2\beta_M^2\psi_{\max}^4-3\psi_{\max}^2-2/3-\gamma^2\left\|I-\beta(k+1)\hat{\psi}(k)\hat{\psi}^T(k)\right\|^2$$
$$D=\gamma^2\left\|I-\beta(k+1)\hat{\psi}(k)\hat{\psi}^T(k)\right\|^2W_{\max}^2+2\beta_M^2\psi_{\max}^2v_M^2+3\mu_M^2\left(1+2\beta_M^2\psi_{\max}^2\right)$$

If the design parameters are selected such that $B_1 > 0$ and $B_2 > 0$, the state and weight estimation errors will be UUB in the mean with the following bounds given by

$$E\left(\left\|\tilde{x}(k)\right\|^2\right) \le \frac{D}{B_1} \quad , \quad E\left(\left\|\tilde{W}(k)\right\|^2\right) \le \frac{D}{B_2}$$

Remark 4: The conditions in Theorem 1 can be satisfied by proper selection of the basis function ψ and user defined parameters including β_M, ψ_{max}, c_ψ, and γ.

In summary, the proposed NN and weight update law can both detect and remove outliers from the measured data or system state or output vector y. First of all, if $\beta(k)$ is zero then an outlier is detected in $y(k)$ which will have no impact on NN weight update. Therefore, outliers will automatically be removed in the state estimates x_s and measurement noise will be moderated. Further, by reducing the effect of outliers on the weight update law, the state estimation issue is resolved and boundedness of state and parameter estimation errors can be obtained similar to the case of no outliers.

After detection and removal of outliers, the estimated outlier-free state vector x_s can be used for fault detection without the risk of having false alarms. Figure 3.1 shows an overview of the combined online outlier detection/removal and fault detection scheme. The next section briefly discusses fault diagnosis after outlier removal.

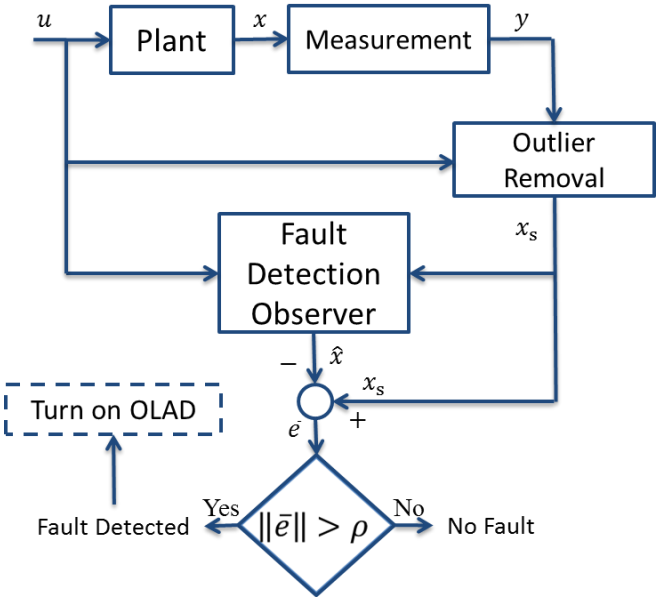

Figure 3.1. Overview of the combined outlier removal and fault detection scheme

4. FAULT DIAGNOSIS SCHEME

Model-based fault detection schemes require an observer to estimate the system states. Then a fault detection (FD) residual will be generated by comparing the actual and estimated system states [28]. Traditional model-based fault detection schemes use the measured states in the fault detection observer and then compare them with observer states to detect faults. But when the measured states are not reliable and involve outliers, false alarms could be triggered. Therefore in this section the outlier-free state vector x_s is used for the purpose of fault diagnosis, instead of the actual measured state vector y. As mentioned in previous section, it can be shown that $\tilde{x} = x - x_s$ is bounded in the mean, i.e. $E(\|\tilde{x}(k)\|) < \sigma_M = \sqrt{D/B_1}$.

Consider the nonlinear FD estimator

$$\hat{x}(k+1) = A_d\hat{x}(k) + \omega\left(x_s(k), u(k)\right) + \hat{h}_d\left(x_s(k), u(k); \hat{\theta}_d(k)\right) - A_d x_s(k) \quad (5)$$

where $\hat{x}(k) \in \mathbb{R}^n$ is the estimated state vector, $\hat{h}_d: \mathbb{R}^n \times \mathbb{R}^{p\times n} \to \mathbb{R}^n$ is the output of the online approximator in discrete time (OLAD) with $\hat{\theta}_d \in \mathbb{R}^{p\times n}$ being its set of unknown parameters, and A_d is a user defined diagonal matrix, which must be selected in a way that the eigenvalues of the closed loop system lie within the unit circle [27]. Initial values of the FD estimator are taken to be $\hat{x}(0) = \hat{x}_0$, $\hat{\theta}_d(0) = \hat{\theta}_{d_0}$, such that $\hat{h}\left(x_s, u, \hat{\theta}_{d_0}\right) = 0$.

In the proposed FD estimator, NNs are used as the OLADs. NN-based OLAD is off prior to the detection of a fault and thus its output is zero. Upon detection of a fault the OLAD is turned on to estimate the fault dynamics.

Define the detection residual as $\bar{e} = x_s - \hat{x}$. Prior to the detection of a fault, the residual dynamics are given by

$$\begin{aligned}
\bar{e}(k+1) &= x_s(k+1) - \hat{x}(k+1) \\
&= x(k+1) - \tilde{x}(k+1) - \hat{x}(k+1) \\
&= A_d\bar{e}(k) + \omega\left(x(k), u(k)\right) - \omega\left(x_s(k), u(k)\right) + \eta\left(x(k), u(k)\right) - \tilde{x}(k+1) \\
&= A_d\bar{e}(k) + \tilde{\omega}(k) + \eta\left(x(k), u(k)\right) - \tilde{x}(k+1)
\end{aligned}$$

where $\tilde{\omega}(k) = \omega\left(x(k), u(k)\right) - \omega\left(x_s(k), u(k)\right)$. As mentioned earlier $\tilde{x}$ which is the difference between actual system states x and estimated outlier-free states x_s, is bounded. Also from assumptions 1 and 2, we know that $\tilde{\omega}$ and η are bounded. Therefore with the

appropriate selection of A_d, the detection residual $\bar{e}$ will remain bounded in healthy operating conditions of the system.

Now consider a dead-zone operator

$$D[z]=\begin{cases}0, & if\ |z|\le\rho\\ z, & if\ |z|>\rho\end{cases}$$

where ρ is the FD threshold. When the detection residual exceeds the detection threshold, a fault is declared active through the dead-zone operator and the OLAD that generates $\hat{h}_d(k)$ is initiated and tuned online using the following update law

$$\hat{\theta}_d(k+1)=\hat{\theta}_d(k)+\alpha\hat{\phi}(k)D\left[\bar{e}^T(k)\right]-\gamma\left\|I-\alpha\hat{\phi}(k)\hat{\phi}^T(k)\right\|\hat{\theta}_d(k) \tag{6}$$

where $\alpha>0$ is the learning rate, $0<\gamma<1$ is the forgetting factor, and $\hat{\phi}(k)=\phi\left(x_s(k),u(k)\right)$ is a basis function like sigmoid or RBF. Moreover the output of the OLAD will be given by

$$\hat{h}_d(k)=\hat{\theta}_d^T(k)\phi\left(x_s(k),u(k)\right)$$

After detection, the residual dynamics can be described by

$$\begin{aligned}\bar{e}(k+1)=A_d\bar{e}(k)+\tilde{\omega}(k)+\eta\left(x(k),u(k)\right)-\tilde{x}(k+1)\\ +h\left(x(k),u(k)\right)-\hat{h}_d\left(x_s(k),u(k);\hat{\theta}_d(k)\right)\end{aligned}$$

Asserting the NLIP assumption on the fault function the above equation can be rewritten as

$$\begin{aligned}\bar{e}(k+1)=A_d\bar{e}(k)+\tilde{\omega}(k)+\eta\left(x(k),u(k)\right)\\ -\tilde{x}(k+1)+\tilde{\theta}_d^T(k)\phi\left(x_s(k),u(k)\right)+\varepsilon(k)\end{aligned} \tag{7}$$

where $\tilde{\theta}_d(k)=\theta_d-\hat{\theta}_d(k)$ is the OLAD parameter estimation error, and $\varepsilon(k)$ is the approximation error which is bounded by ε_M.

The stability of the proposed scheme will be investigated in the following theorem

Theorem 2 (Fault Diagnosis Observer Performance): Let the proposed observer in (5) be used to monitor the system in (1), with the OLAD turned on upon detection of a fault. Let the update law in (6) be used to update the unknown parameter set $\hat{\theta}_d$. Then the FD residual, $\bar{e}(k)$, and the parameter estimation errors, $\tilde{\theta}_d(k)$ are uniformly ultimately bounded in the mean.

Proof : Consider the following Lyapunov function candidate

$$V = E\left(tr\left\{\bar{e}^T(k)\bar{e}(k)\right\}\right) + E\left(tr\{\tilde{\theta}_d^T(k)\tilde{\theta}_d(k)\}\right)$$

Then the first difference of the Lyapunov function is given by

$$\Delta V = \underbrace{E\left(tr\left\{\bar{e}^T(k+1)\bar{e}(k+1) - \bar{e}^T(k)\bar{e}(k)\right\}\right)}_{\Delta V_1} + \underbrace{E\left(tr\left\{\tilde{\theta}_d^T(k+1)\tilde{\theta}_d(k+1) - \tilde{\theta}_d^T(k)\tilde{\theta}_d(k)\right\}\right)}_{\Delta V_2}$$

By substituting $\bar{e}(k+1)$ from (7) in ΔV_1 and applying the Cauchy-Schwarz inequality we get

$$\begin{aligned}\Delta V_1 \le E\Big(tr\Big\{&6\bar{e}^T(k)A_d^T A_d\bar{e}(k) + 6\tilde{\omega}^T(k)\tilde{\omega}(k) + 6\eta^T\left(x(k),u(k)\right)\eta\left(x(k),u(k)\right)\\ &+ 6\tilde{x}^T(k+1)\tilde{x}(k+1) + 6\hat{\phi}^T(k)\tilde{\theta}_d(k)\tilde{\theta}_d^T(k)\hat{\phi}(k) + 6\varepsilon^T(k)\varepsilon(k) - \bar{e}^T(k)\bar{e}(k)\Big\}\Big)\\ \le\ & 6E\left(tr\left\{A_d^T A_d\bar{e}(k)\bar{e}^T(k)\right\}\right) + 6E\left(\tilde{\omega}^T(k)\tilde{\omega}(k)\right)\\ &+ 6E\left(\eta^T\left(x(k),u(k)\right)\eta\left(x(k),u(k)\right)\right)\\ &+ 6E\left(\tilde{x}^T(k+1)\tilde{x}(k+1)\right) + 6E\left(\varepsilon^T(k)\varepsilon(k)\right)\\ &+ 6E\left(tr\left\{\hat{\phi}(k)\hat{\phi}^T(k)\tilde{\theta}_d(k)\tilde{\theta}_d^T(k)\right\}\right) - E\left(\bar{e}^T(k)\bar{e}(k)\right)\end{aligned}$$

Now substitute $\tilde{\theta}_d(k+1)$ from (6) in ΔV_2 and use the Cauchy-Schwarz inequality to arrive at

$$\begin{aligned}\Delta V_2 = E\Big(tr\Big\{&\Big(\tilde{\theta}_d(k) - \alpha\hat{\phi}(k)\bar{e}^T(k) - \gamma\left\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\right\|\tilde{\theta}_d(k)\\ &+ \gamma\left\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\right\|\theta_d\Big)^T\Big(\tilde{\theta}_d(k) - \alpha\hat{\phi}(k)\bar{e}^T(k)\\ &- \gamma\left\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\right\|\tilde{\theta}_d(k) + \gamma\left\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\right\|\theta_d\Big) - \tilde{\theta}_d^T(k)\tilde{\theta}_d(k)\Big\}\Big)\\ \le\ & 3\alpha^2\hat{\phi}^T(k)\hat{\phi}(k)E\left(\bar{e}^T(k)\bar{e}(k)\right) + 3\gamma^2\left\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\right\|^2 E\left(tr\left\{\theta_d^T\theta_d\right\}\right)\\ &+ 2E\left(tr\left\{\tilde{\theta}_d^T(k)\tilde{\theta}_d(k)\right\}\right) - 6\gamma\left\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\right\|E\left(tr\left\{\tilde{\theta}_d^T(k)\tilde{\theta}_d(k)\right\}\right)\\ &+ 3\gamma^2\left\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\right\|^2 E\left(tr\left\{\tilde{\theta}_d^T(k)\tilde{\theta}_d(k)\right\}\right)\end{aligned}$$

By combining ΔV_1 and ΔV_2, taking Frobenius norm, and using assumption 2, we arrive at

$$\Delta V \le 6A_{d_{max}}^2 E\left(\left\|\bar{e}(k)\right\|^2\right) + 6c_g^2 E\left(\left\|\tilde{x}(k)\right\|^2\right) + 6\eta_M^2 + 6\sigma_M^2$$
$$+ 6\varepsilon_M^2 + 6\phi_{max}^2 E\left(\left\|\tilde{\theta}_d(k)\right\|^2\right) - E\left(\left\|\bar{e}(k)\right\|^2\right)$$
$$+ 3\alpha^2\phi_{max}^2 E\left(\left\|\bar{e}(k)\right\|^2\right) + 3\gamma^2\left\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\right\|^2 \theta_{d_{\max}}^2$$

$$\Delta V \le 6A_{d_{max}}^2 E\left(\left\|\bar{e}(k)\right\|^2\right) + 6c_g^2 E\left(\left\|\tilde{x}(k)\right\|^2\right) + 6\eta_M^2 + 6\sigma_M^2$$
$$+ 6\varepsilon_M^2 + 6\phi_{max}^2 E\left(\left\|\tilde{\theta}_d(k)\right\|^2\right) - E\left(\left\|\bar{e}(k)\right\|^2\right)$$
$$+ 3\alpha^2\phi_{max}^2 E\left(\left\|\bar{e}(k)\right\|^2\right) + 3\gamma^2\left\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\right\|^2 \theta_{d_{\max}}^2$$
$$- 6\gamma\left\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\right\| E\left(\left\|\tilde{\theta}_d(k)\right\|^2\right)$$
$$+ \left(3\gamma^2\left\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\right\|^2 + 2\right) E\left(\left\|\tilde{\theta}_d(k)\right\|^2\right)$$

$$\le -\left(1 - 6A_{d_{max}}^2 - 3\alpha^2\phi_{max}^2\right) E\left(\left\|\bar{e}(k)\right\|^2\right)$$
$$- \left(6\gamma\left\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\right\| - 6\phi_{max}^2\right.$$
$$\left. - 3\gamma^2\left\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\right\|^2 - 2\right) E\left(\left\|\tilde{\theta}_d(k)\right\|^2\right)$$
$$+ 6c_g^2\sigma_M^2 + 6\eta_M^2 + 6\sigma_M^2 + 6\varepsilon_M^2 + 3\gamma^2\left\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\right\|^2 \theta_{d_{\max}}^2$$

Therefore, $\bar{e}$ and $\tilde{\theta}_d$ are uniformly ultimately bounded in the mean if the following conditions are satisfied

$$6A_{d_{max}}^2 + 3\alpha^2\phi_{max}^2 < 1$$

$$3\gamma\left\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\right\|\left(2 - \gamma\left\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\right\|\right) > 6\phi_{max}^2 + 2$$

Moreover, the bounds are given by

$$E\left(\left\|\bar{e}(k)\right\|^2\right) \le \frac{F}{C_1} \quad , \quad E\left(\left\|\tilde{\theta}_d(k)\right\|^2\right) \le \frac{F}{C_2}$$

where F, C_1, and C_2 are defined by

$$F = 6c_g^2\sigma_M^2 + 6\eta_M^2 + 6\sigma_M^2 + 6\varepsilon_M^2 + 3\gamma^2\left\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\right\|^2 \theta_{d_{\max}}^2$$

$$C_1 = 1 - 6A_{d_{max}}^2 - 3\alpha^2\phi_{max}^2$$

$$C_2 = 6\gamma\left\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\right\| - 6\phi_{max}^2 - 2 - 3\gamma^2\left\|I - \alpha\hat{\phi}(k)\hat{\phi}^T(k)\right\|^2$$

5. SIMULATION RESULTS

In this section a three tank water system is selected to verify the performance of the proposed schemes in simulations and then an axial piston pump testbed is used as an experimental study to show the effectiveness of the proposed outlier removal technique in practice.

5.1. EXAMPLE 1: THREE-TANK BENCHMARKING SYSTEM

A schematic view of the three-tank benchmarking system [29] is shown Figure 5.1. This system consists of three tanks connected to each other, two input pumps on tank 1 and tank 2 and one water outlet on tank 2.

The three-tank system dynamics are described by $x(k+1)=\omega\big(x(k),u(k)\big)+\eta\big(x(k)\big)$ where $x=[x_1,x_2,x_3]^T$ is the state vector and $\omega\big(x(k),u(k)\big)$ is the known nonlinear dynamics of the system [20] given by

$$\omega\big(x(k),u(k)\big)=\begin{bmatrix} \frac{T}{A}\{-c_1S_p sign\big(x_1(k)-x_3(k)\big)\sqrt{2g|x_1(k)-x_3(k)|} \\ +u_1(k)\}+x_1(k) \\ \frac{T}{A}\{-c_3S_p sign\big(x_2(k)-x_3(k)\big)\sqrt{2g|x_2(k)-x_3(k)|} \\ -c_2S_p\sqrt{2gx_2(k)}+u_2(k)\}+x_2(k) \\ \frac{T}{A}\{-c_1S_p sign\big(x_1(k)-x_3(k)\big)\sqrt{2g|x_1(k)-x_3(k)|} \\ -c_3S_p sign\big(x_3(k)-x_2(k)\big)\sqrt{2g|x_3(k)-x_2(k)|}\}+x_3(k) \end{bmatrix}$$

where T is the sampling time chosen to be 0.01 seconds, $A=0.0154\,m^2$ is the cross section of the tanks, $S_p=5\times10^{-5}\,m^2$ is the cross section of the connecting pipes, $c_1=1, c_2=0.8,$ and $c_3=1$ are the outflow coefficients, and $g=9.8\,m/s^2$ is the standard gravity. Moreover

$$\eta\big(x(k)\big)=[10^{-3}\sin(0.7kT)\ \ 10^{-2}\cos(0.8kT)\ \ 10^{-1.65}\cos(0.5kT)]^T$$

represents the modeling uncertainty.

This system is subjected to a fault which is given in terms of leakage in tank 1 and occurs at time t_0 =40 s. The fault function is described by

$$h\big(x(k)\big)=\left[0.0154\left(1-e^{-0.5T(k-k_0)}\right)\sqrt{2gx_1(k)},\ 0,\ 0\right]^T$$

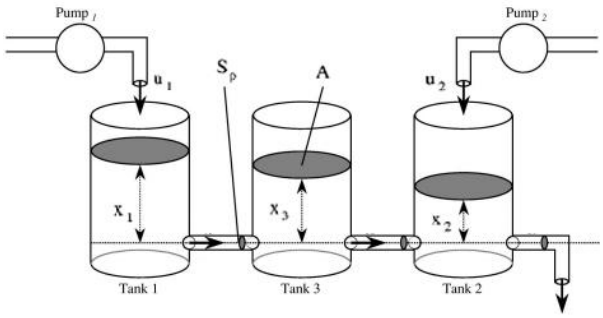

Figure 5.1. Schematic view of the three-tank system

The FD estimator in (5) is used to detect the faults, where $A_d = 0.99I_{3\times3}$. The OLAD output is given by $\hat{h}_d(k) = \hat{\theta}_d^T(k)\phi(Vx_s(k) + B)$, where $\hat{\theta}_d \in \mathbb{R}^{8\times3}$ is the estimated parameters while $\phi \in \mathbb{R}^8$ is a vector of sigmoid functions. Moreover V, B are selected randomly and the update law parameters are $\alpha = 0.5, \gamma = 10^{-4}$. The detection threshold, ρ, is selected to be 2.5.

Figure 5.2 shows the actual system states while Figure 5.3 depicts the measured system states involving a number of outliers. The state distribution for healthy and faulty periods is depicted in Figure 5.4. The mean and variance of the distribution in healthy period are 2.33 and 2.05 respectively, while mean and variance values for the faulty period are 54.24 and 1479.53, respectively. If the measured data is used for fault detection, the outliers will cause false alarms to be triggered. This can be observed in Figure 5.5 where the detection residual is plotted along with detection threshold. Fault is seeded at t=40 s, but a false alarm will be triggered at about t=8 s. It is obvious that in this case, the estimated fault given by the OLAD cannot be close to the actual fault at all.

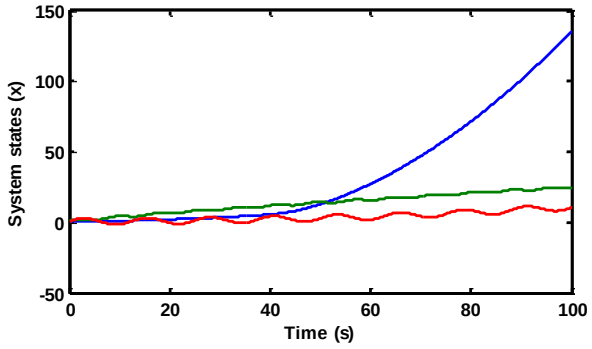

Figure 5.2. Actual system states x

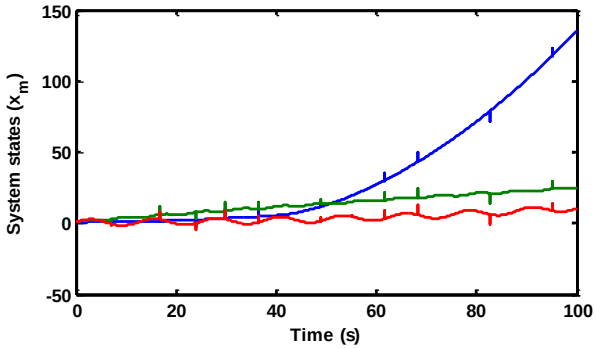

Figure 5.3. Measured system states

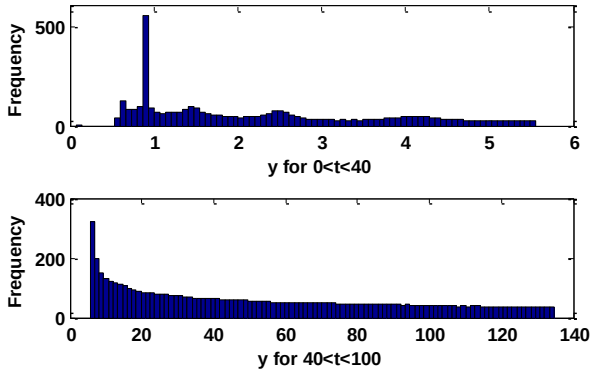

Figure 5.4. Distribution of the measured data y_1

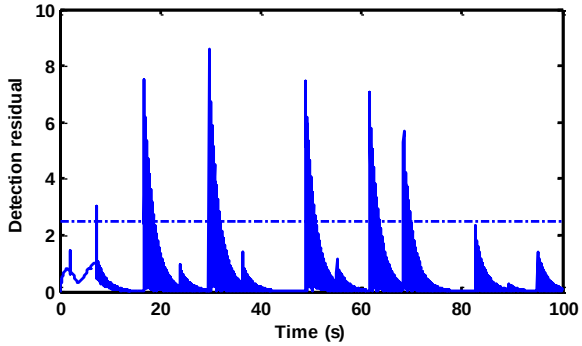

Figure 5.5. Detection residual without outlier removal

In order to fix this problem, the proposed outlier removal scheme is first utilized to remove outliers from the measured data. A NN with 12 hidden layer neurons and sigmoid activation functions is used to estimate x_s and p which is the window size is selected as 100 (that means 1 second). The estimated x_s is shown in Figure 5.6. It can be

observed that the outliers are removed from the measured states. Table 5.1 shows the outlier detection results for several cases with different number of outliers at randomly selected times. It is observed that the proposed scheme has been able to detect 100% of the outliers in most cases, with low number of false positives.

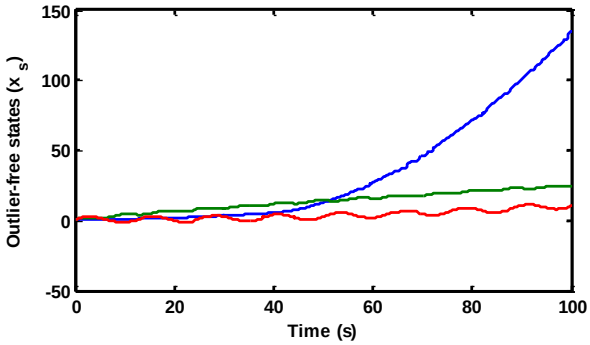

Figure 5.6. Estimated outlier-free states x_s

Table 5.1. Outlier detection statistics

* The presented plots correspond to case 1

Case #	Total number of outliers	Number of True positives	Number of False positives
1*	12	12	1
2	8	8	0
3	14	14	2
4	18	17	1
5	9	9	0
6	11	11	0

When x_s is used in the fault detection observer, no false alarm will be triggered and the actual fault is detected at t=74 seconds (Figure 5.7). Figure 5.8 shows the estimated and actual fault magnitudes in this case. Unlike the previous case, when no outlier removal was performed, the fault can be estimated with a small error. The simulation results clearly illustrate the effectiveness of the proposed outlier

detection/removal scheme. Furthermore the importance of removing the outliers before performing fault detection is clarified.

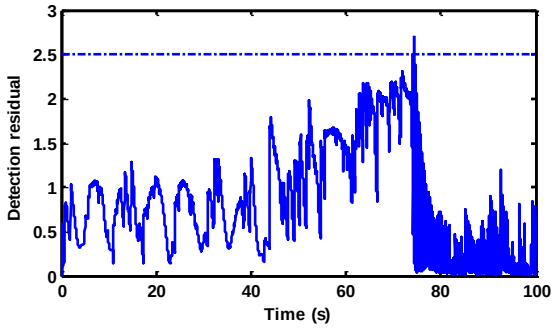

Figure 5.7. Detection residual with outlier removal

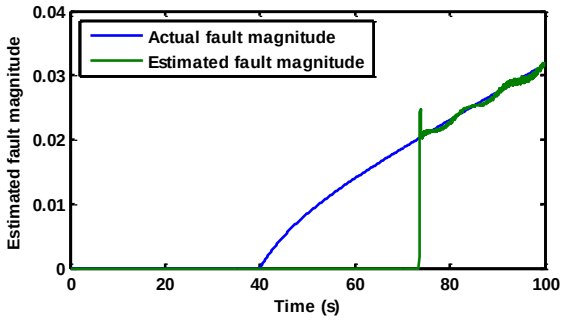

Figure 5.8. Detection residual with outlier removal

5.2. EXAMPLE 2: LINEAR MATHEMATICAL EXAMPLE

A linear system has been selected as the second example to compare the proposed scheme with a Kalman filter-based approach. The system is described by $x(k+1) = Ax(k)$, where A is defined by

$$A = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \quad for\ t \le 50$$

$$A = 1.003\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \quad for\ t > 50$$

with $\theta = \pi/500$. In fact the system dynamics are slightly changed during the simulation, in order to test the proposed scheme and compare its performance with a Kalman filter-based method under nonstationary operating conditions. The entire simulation time is

taken to be 100 seconds and the sampling time is 0.1 seconds, and the states are measured as follows $y(k) \sim \mathcal{N}\left(x(k), \begin{bmatrix} .25 & 0 \\ 0 & 0.25 \end{bmatrix} + \begin{bmatrix} r_1(k) & 0 \\ 0 & r_2(k) \end{bmatrix} \right)$ where $r_1(k)$ is set to 1 with probability κ and to zero with probability $1 - \kappa$.

Figure 5.9 shows the actual system states x, along with measured states y when $\kappa = 0.05$ and Figure 5.10 depicts the distribution of $y_2 - x_2$ for $t < 50$ and $t > 50$. In the first half of the simulation the mean and variance of $y_2 - x_2$ are 0.049 and 0.82 respectively, whereas mean and variance in the second half of the simulation are -0.008 and 1.37. The proposed outlier detection and removal scheme is applied on the measured data, with $p = 10$, $\gamma = 0.01$, and randomly selected W matrix. The estimated states are shown in Figure 5.11. It is worth mentioning that the performance of the scheme is not degraded after the change in the system dynamics at t=50 s.

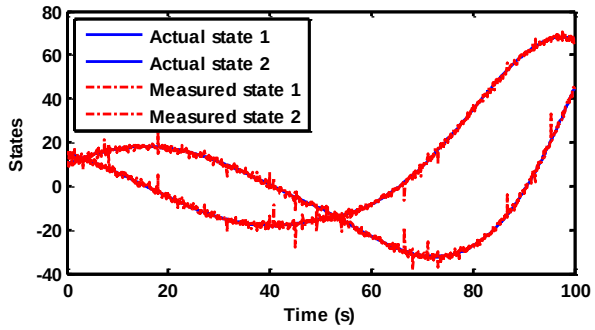

Figure 5.9. Actual and measured system states

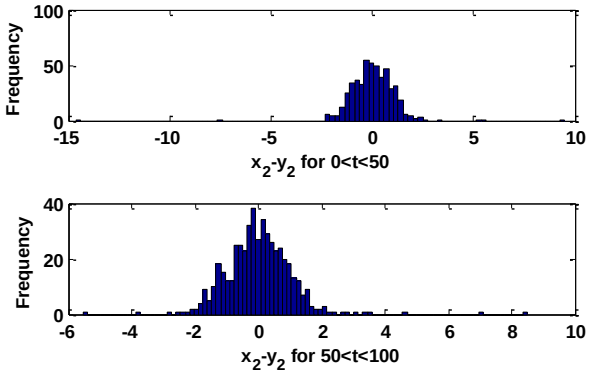

Figure 5.10. Distribution of the difference between actual and measured states

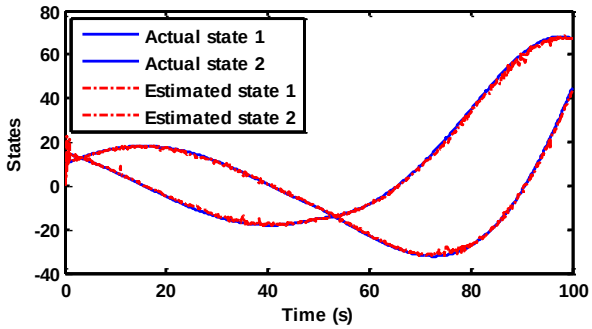

Figure 5.11. Estimated states using the proposed NN-based scheme

Next an outlier-robust Kalman filter whose parameters are fitted according to the Maximum Likelihood criterion [19], is applied on the same measured data and the result is shown in Figure 5.12. Although this method has a good performance in the first 50 seconds of the simulation, its performance is extremely degraded when the dynamics of the system changes. Mean squared error of state estimation for both of the methods are presented in Table 5.2 for comparison. This simulation clearly shows that unlike Kalman filter-based approaches, our outlier removal method is robust to changes in the system dynamics (which could be due to faults or changes in the operating conditions). Although the state estimation error of the Kalman filter method after the occurrence of fault seems to be useful for fault detection, the large error in estimation makes this method useless for outlier removal in the presence of fault.

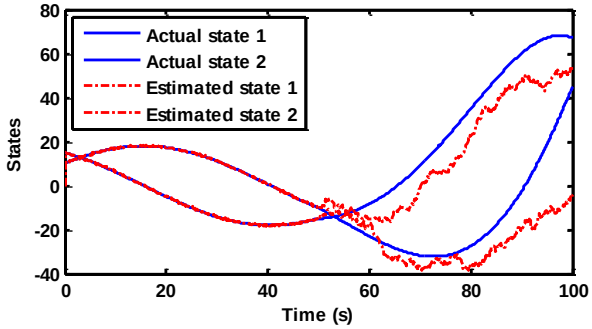

Figure 5.12. Estimated states using an outlier robust Kalman filter

Table 5.2. Outlier removal performance comparison

Method	Mean squared error
Proposed scheme	2.14
Outlier robust Kalman filter	110.29

The simulations have been repeated for 4000 times using the proposed method. The noise and outliers are random, thus vary from one simulation to the other. The average and maximum mean squared error of state estimation and the average percentage of detected outliers are shown in Table 5.3. The results imply that the proposed method is able to detect and remove the outliers with consistently high performance. The important point is that the average percentage of detected outliers is as high as 97% and the maximum mean squared error in all the simulations is less than 4 which is still very low compared to the Kalman filter method [19]**.**

Table 5.3. Results of the repeated simulations using NN-based scheme

Total number of simulations	Average MSE	Maximum MSE	Average percentage of detected outliers
4000	1.98	3.56	%97

5.3. EXAMPLE 3: AXIAL PISTON PUMP

An axial piston pump testbed is used to test the performance of the proposed outlier removal scheme and observe its effect on fault detection in an experimental study. A picture of this testbed is shown in Figure 5.13. The nonlinear dynamics of this system is described by

$$x_i(k+1) = x_i(k) + \frac{BT}{C - A_p S_{pi}}\left(Q_{kpi}(k) - Q_{pi}(k) - Q_{lpi}(k)\right) \quad for\ i = 1,...,9$$

$$x_{10}(k+1) = x_{10}(k) + \frac{CBT}{V_c}\left(\sum_{i=1}^{9} Q_{pi}(k) - C_{d2} A_v \sqrt{\frac{2x_{10}(k)}{\rho_c}}\right)$$

where $[x_1(k), \ldots, x_9(k)]$ represent the pressures in the nine pistons, $x_{10}(k)$ is the pump outlet pressure, B is the bulk modulus of the hydraulic fluid, V_c is the theoretical volume of flow, A_p is the piston area, C_{d2} is the discharge coefficient of needle valve orifice, $A_v$is the orifice area of the needle valve, ρ_c is the flow density, and T is the sampling time. Furthermore $Q_{kpi}(k), Q_{lpi}(k), Q_{pi}(k), S_{pi}(k)$ are obtained by

$$Q_{kpi}(k) = \frac{\omega\pi d^2 R_p}{4}\tan(\beta_c)\sin(\omega k - (i-1)\alpha_p), Q_{lpi}(k) = \frac{\pi r h_g^3}{6\mu L}(x_i(k) - P_c)$$

$$Q_{pi}(k) = C_{d1}A_{d1}\sqrt{\frac{2|x_i(k) - x_{10}(k)|}{\rho_c}}$$

$$S_{pi}(k) = R_p\tan(\beta_c)\left(1 - \cos(\omega k - (i-1)\alpha_c)\right)$$

where ω is the angular velocity of the pump drive shaft, d is the diameter of the piston, R_p is the piston radius on barrel, $\beta_c$is the angle of swash plate, α_p is the phase delay, r is the radius of piston, μ is the absolute fluid velocity, and L is the length of leakage passage.

Figure 5.13. Picture of the axial piston pump testbed

In this system only one of the states, namely the pump outlet pressure is measurable. Therefore for the purpose of fault detection an output observer [24] is constructed using the model of the system. The dynamics of the output observer is slightly different from the full state observer presented in this paper, in that it uses the

output of the system instead of the entire state vector. The sampling time for measuring the data is 0.1 second. The output of the system is measured in healthy operating conditions for 200 seconds and is shown in Figure 5.14. Mean and variance of the whole data set are 1428.79 and 59.34 respectively.

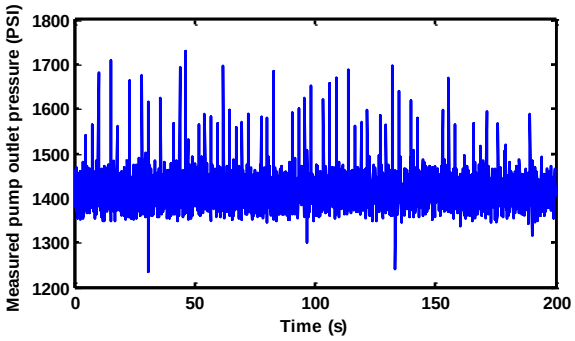

Figure 5.14. Measured pump outlet pressure

If the measured data is directly used for fault detection, several false alarms will be triggered. This can be clearly observed in Figure 5.15 which shows the fault detection residual and threshold without any outlier removal performed. To solve this problem, we use our proposed outlier removal scheme. The user defined parameters of the update law are selected as $\tau = 0.2$ and $p = 20$. The estimated state x_s is plotted in Figure 5.16 and the fault detection residual when x_s is used for fault detection is shown in Figure 5.17. It is clearly observed that the outliers are removed and no false alarm is triggered.

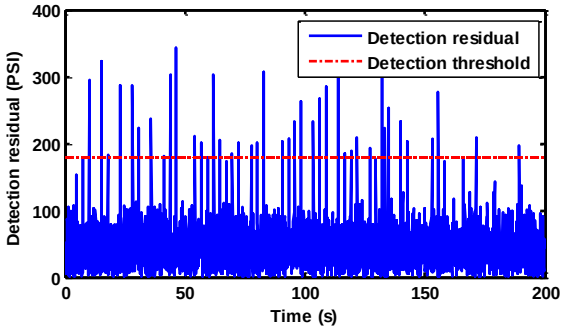

Figure 5.15. Detection residual without outlier removal

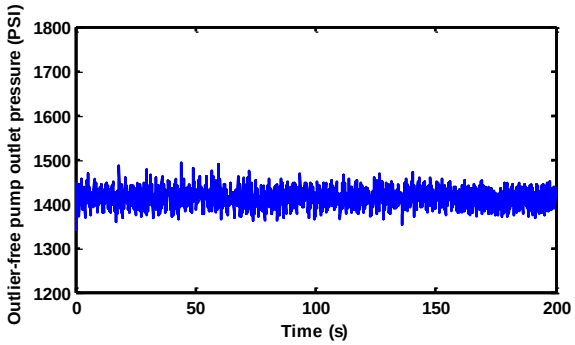

Figure 5.16. Estimated outlier-free pump outlet pressure

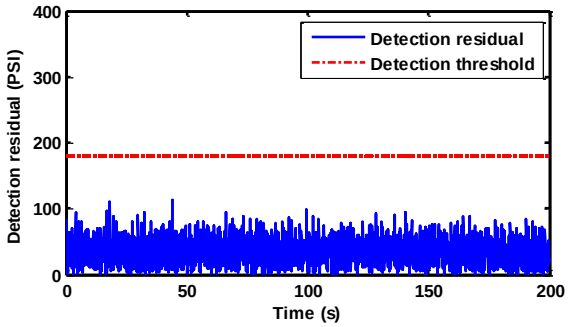

Figure 5.17. Detection residual with outlier removal

6. CONCLUSIONS

In this paper, a NN-based online outlier detection and removal scheme was presented and combined with a model-based fault detection scheme. It was demonstrated that the underlying distribution of data in the case of a data driven scheme or states in the case of model-based fault detection is nonstationary due to presence of changing dynamics, outliers and noise. Moreover, it was shown that a separate OIR scheme is necessary prior to any fault detection and diagnosis. On the other hand when the outliers are removed by the proposed scheme, fault detection can be performed successfully. The proposed observer-based method changes the learning rate to zero when an outlier is detected. The proposed OIR scheme can function even when the measured data which is going to be used for monitoring and fault detection is contaminated with outliers. Then a data driven fault detection scheme can yield low detection rate and high false alarm rate similar to the model-based fault detection framework.

7. REFERENCES

[1] D. M. Hawkins, *Identification of Outliers*. London – New York: Chapman and Hall, 1980.

[2] S. M. Namburu, M. Wilcutts, S. Chigusa, Q. Liu, C. Kihoon, and K. Pattipati, "Systematic Data-Driven Approach to Real-Time Fault Detection and Diagnosis in Automotive Engines," *IEEE Autotestcon*, pp. 59-65, 2006.

[3] R. Isermann, "Model-based fault-detection and diagnosis – status and applications," *Annual Reviews in Control,* vol. 29, pp. 71-85, 2005.

[4] B. J. Ohran, J. Liu, D. M. de la Peña, P. D. Christofides, and J. F. Davis, "Data-based fault detection and isolation using feedback control: Output feedback and optimality," *Chemical Engineering Science,* vol. 64, pp. 2370-2383, 2009.

[5] R. Ghimire, C. Sankavaram, A. Ghahari, K. Pattipati, Y. Ghoneim, M. Howell, and M. Salman, "Integrated model-based and data-driven fault detection and diagnosis approach for an automotive electric power steering system," *IEEE Autotestcon*, pp. 70-77, 2011.

[6] X. Zhang, M. M. Polycarpou, and T. Parisini, "Fault diagnosis of a class of nonlinear uncertain systems with Lipschitz nonlinearities using adaptive estimation," *Automatica,* vol. 46, pp. 290-299, 2010.

[7] D. Blake and M. Brown, "Simultaneous, Multiplicative Actuator and Sensor Fault Estimation using Fuzzy Observers," *IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2007*, pp. 1-6, 2007.

[8] H. A. Talebi and K. Khorasani, "A neural network-based actuator gain Fault Detection and Isolation strategy for nonlinear systems," *46th IEEE Conference on Decision and Control*, pp. 2614-2619, 2007.

[9] G. Jiaxing and T. Gang, "A feedback-based fault detection scheme for aircraft systems with damage," *8th Asian Control Conference (ASCC)*, pp. 1431-1436, 2011.

[10] H. Ferdowsi and S. Jagannathan, "A unified model-based fault diagnosis scheme for non-linear discrete-time systems with additive and multiplicative faults," *Transactions of the Institute of Measurement and Control*, vol. 35, pp. 742-752, 2013.

[11] P. J. Rousseeuw and K. V. Driessen, "A fast algorithm for the minimum covariance determinant estimator," *Technometrics,* vol. 41, pp. 212-223, 1999.

[12] E. M. Knorr, R. T. Ng, and V. Tucakov, "Distance-based outliers: algorithms and applications," *The VLDB Journal*, vol. 8, pp. 237-253, 2000.

[13] E. M. Knorr and R. T. Ng, "Algorithms for Mining Distance-Based Outliers in Large Datasets," *Proceedings of the 24rd International Conference on Very Large Data Bases*, pp. 392-403, 1998.

[14] G. Kollios, D. Gunopulos, N. Koudas, and S. Berchtold, "Efficient biased sampling for approximate clustering and outlier detection in large data sets," *IEEE Transactions on Knowledge and Data Engineering*, vol. 15, pp. 1170-1187, 2003.

[15] Z. Jiang-She and L. Yiu-Wing, "Robust clustering by pruning outliers," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 33, pp. 983-998, 2003.

[16] M. M. Breunig, H.-P. Kriegel, R. T. Ng, and R. Sander, "LOF: identifying density-based local outliers," *Proceedings of the 2000 ACM SIGMOD international conference on Management of data*, pp. 93-104, 2000.

[17] V. J. Hodge and J. Austin, "A survey of outlier detection methodologies," *Artificial Intelligence Review*, vol. 22, pp. 85-126, 2004.

[18] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Computing Surveys (CSUR)*, vol. 41, p. 15, 2009.

[19] G. Agamennoni, J. I. Nieto, and E. M. Nebot, "An outlier-robust Kalman filter," *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1551-1558, 2011.

[20] S. C. Chan, Z. G. Zhang, and K. W. Tse, "A new robust Kalman filter algorithm under outliers and system uncertainties," *IEEE International Symposium on Circuits and Systems, ISCAS 2005*, vol. 5, pp. 4317-4320, 2005.

[21] J. A. Ting, E. Theodorou, and S. Schaal, "A Kalman filter for robust outlier detection," *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2007*, pp. 1514-1519, 2007.

[22] J. J. Gertler, "Survey of model-based failure detection and isolation in complex plants," *IEEE Control Systems Magazine*, vol. 8, pp. 3-11, 1988.

[23] N. Tanabe, T. Furukawa, H. Matsue, and S. Tsujii, "Kalman filter for robust noise suppression in white and colored noises," *IEEE International Symposium on Circuits and Systems, ISCAS 2008*, pp. 1172-1175, 2008.

[24] B. T. Thumati and S. Jagannathan, "A Model-Based Fault-Detection and Prediction Scheme for Nonlinear Multivariable Discrete-Time Systems with Asymptotic Stability Guarantees," *IEEE Transactions on Neural Networks*, vol. 21, pp. 404-423, 2010.

[25] B. G. Amidan, T. A. Ferryman, and S. K. Cooley, "Data outlier detection using the Chebyshev theorem," *IEEE Aerospace Conference*, pp. 3814-3819, 2005.

[26] Z. Yue, L. Jie, and S. Bo, "A New Algorithm for Outlier Detection Based on Offset," *Fifth International Conference on Information Assurance and Security, IAS '09.*, pp. 3-6, 2009.

[27] S. Jagannathan, *Neural Network Control of Nonlinear Discrete –time Systems*. NY: CRC publications, 2006.

[28] F. Caccavale, P. Cilibrizzi, F. Pierri, and L. Villani, "Actuators fault diagnosis for robot manipulators with uncertain model," *Control Engineering Practice,* vol. 17, pp. 146-157, 2009.

[29] D. Theilliol, H. Noura, and J.-C. Ponsart, "Fault diagnosis and accommodation of a three-tank system based on analytical redundancy," *ISA Transactions,* vol. 41, pp. 365-382, 2002.

III. Decentralized Fault Diagnosis and Prognosis Scheme for Interconnected Nonlinear Discrete-Time Systems

Hasan Ferdowsi and S. Jagannathan

Abstract

This paper deals with the design of a decentralized fault diagnosis and prognosis methodology for large-scale interconnected nonlinear dynamical discrete-time systems which are modelled as the interconnection of several subsystems. For each subsystem, a local fault detector (LFD) is designed based on dynamic model of the local subsystem and estimated states. Each LFD consists of an observer with an incorporated online approximator. Online approximators only use local measurements as their inputs and are always turned on and continuously learn the interconnection function as well as possible fault dynamics. A fault is detected by comparing the output of each online approximator with a predefined threshold instead of using the residual. Derivation of robust detection thresholds and fault detectability conditions are also included. Due to interconnection nature of the large scale system, the effect of faults propagate to other subsystems, thus a fault might be detected in more than one subsystem. Upon detection, faults local to the subsystem and from other subsystems are isolated by using a central fault isolation unit which receives detection time information from all LFDs. The proposed scheme also provides the time-to-failure or remaining useful life information by using local measurements. Simulation results provide the effectiveness of the proposed decentralized fault detection scheme.

1. INTRODUCTION

Several practical systems such as the well-known power generation and distribution systems, telecommunication networks, water distribution networks, traffic networks, exhibit complex and spatially distributed dynamics and can be referred to as large scale interconnected systems. With increasing complexity with these systems, there is a high possibility of occurrence of faults. Therefore, suitable fault diagnosis schemes which permit the operation of such large scale interconnected systems reliably at all times are needed. In this paper, a quantitative decentralized fault diagnosis scheme for a large-scale interconnected system in discrete-time and its rigorous analysis are introduced.

Out of the data driven and model-based fault diagnosis schemes, data driven methods [1] need healthy and faulty data from the system, which can be quite expensive to collect, store and process. Model-based fault diagnosis schemes [2], on the other hand, do not require significant quantities of data for development whereas require data to detect faults online. Therefore, a number of researchers have worked on model-based FD schemes, using adaptive estimators or observers [3-5], neural network based observers [6,7], fuzzy observers [8,9], etc, for several practical industrial systems.

In the recent literature, decentralized control of distributed systems [10-12] by using local subsystem states is introduced due to the effort involved in transmitting entire system state vector for a centralized control scheme. In contrast, the fault diagnosis articles [4-8] for such interconnected systems offer centralized FD schemes that require all the states of the system to be measured.

Recently by using overlapping decomposition [13], a large-scale system is decomposed into a set of subsystems which are interconnected by unknown nonlinear functions and distributed fault diagnosis scheme is introduced by assuming the entire state vector is available. On the other hand, decentralized fault diagnosis schemes in [13,14] are introduced for continuous-time systems by assuming that the interconnection functions are known and the entire estimated system state vector is available at each subsystem.

However, for large-scale interconnected systems, it is very expensive and time consuming to gather and process all the measured system states while the measurements can be outdated due to delay in transmission although availability of all the state

information at each subsystem can help in an accurate diagnosis. Additionally, time-to-failure (TTF) information is not included upon detection in all the above schemes [8,13,14].

By contrast, our objective in this paper is to design a network of local fault detectors (LFD) or observers for interconnected nonlinear discrete-time systems so that each LFD monitors a single subsystem by making use of the local information or states in contrast with [13,14]. In addition, partial isolation of faults and TTF will be included upon detection.

Since discrete-time implementation is preferred for hardware implementation [15], in this work, the nonlinear system is modeled in discrete-time along with external disturbances, unmodeled dynamics and interconnection effects. The class of faults considered is allowed to be nonlinear with respect to the state and input, and includes both abrupt and incipient faults. Incipient faults may be difficult to deal with owing to the fact that their small effects on residuals can be hidden as if they are due to the modeling uncertainty. Here, we stress the design of truly decentralized fault diagnosis scheme for incipient faults in discrete-time.

Each local fault detector mainly consists of a nonlinear observer with an incorporated online approximator which is used to estimate the unknown part of the subsystem dynamics, i.e. interconnection term and possible fault function. A local residual signal is generated by comparing the estimated local states from the observer with the measured system states. However, this residual is not used for performing fault detection, whereas it is used to update the unknown parameters of the online approximator. In contrast with other model-based fault detection methods [3-9,13,14], the online approximator is always turned on and the detection is performed by comparing the output of the OLAD with a predefined threshold. This is possible due to the fact that the interconnection term remains bounded as long as the system works in healthy operating conditions and the system states remain bounded in the absence of fault. In addition, a mathematically rigorous approach to the derivation of robust detection thresholds and fault detectability condition is given.

This approximator only uses the local states and learns the interconnection and fault dynamics at each subsystem. It is mathematically shown that although the

interconnection term is a function of nonlocal states, it can be estimated by an online approximator whose inputs are the measured local states at current and next time instant. In order to make this method practical, we have considered the observers one time step behind the actual system.

Upon detection, a fault isolation algorithm is utilized to determine whether or not the fault is local by making use of a central fault isolation unit. The proposed isolation scheme requires minimal transmission of information, as the only information which needs to be transmitted is the detection time from the LFDs. Local faults affects local measurements quicker than non-local faults. Therefore by comparing the detection time at all the subsystems that are made available at the centralized isolation unit, the location of the fault is identified and appropriate action can be taken subsequently.

The accurate approximation of interconnection and fault functions allows a good estimation of the states, thus allowing proper estimation of time-to-failure by comparing the system state estimates against the user defined failure limits [16]. The TTF ensures that the system will not be operated beyond this limit as it is unsafe. In this paper, the TTF is determined by using estimated system states instead of parameter estimates.

Thus the major contributions of this paper include the development of a decentralized fault diagnosis scheme for nonlinear discrete-time systems wherein a LFD only uses local measurements in contrast with [3-9]. Each LFD is designed to detect faults regardless of their location and then determine if the fault is local or nonlocal by using a centralized isolation module. Furthermore TTF estimation is performed upon fault detection whereas such scheme is not available in most of the model-based methods [2-5].

This paper is organized as follows. Section 2 gives a problem formulation for large-scale interconnected nonlinear discrete-time systems. Section 3 proposes the decentralized fault detection scheme including the main results on detectable faults, and discusses the partial isolation of faults as well as TTF determination, and finally Section 4 reports simulation results.

2. SYSTEM DESCRIPTION

Consider the interconnected nonlinear discrete-time systems described by

$$x(k+1)=\omega(x(k),u(k))+\eta(x(k),u(k))+\Pi(k-k_0)h(x(k),u(k))$$

where $u\in\mathbb{R}^m$ is the control input vector, $x\in\mathbb{R}^n$ is the system state vector, $\omega:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$ represents the nonlinear system dynamics, $\eta:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$ represents the system uncertainties, and $h:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^n$ represents a vector of possible fault dynamics. Suppose that this system is comprised of N interconnected subsystems. The ith subsystem dynamics are given by

$$\begin{aligned}x_i(k+1)=&\ f_i(x_i(k),u_i(k))+g_i(x_i(k),\bar{x}_i(k),u_i(k))\\&+\eta_i(x_i(k),u_i(k))+\Pi(k-k_0)h_i(x_i(k),u_i(k))\end{aligned}\tag{1}$$

where $u_i\in\mathbb{R}^{m_i}$ is the local control input vector, $x_i\in\mathbb{R}^{n_i}$ is the local state vector, $\bar{x}_i\in\mathbb{R}^{\bar{n}_i}$ is the vector of interconnection states, $f_i:\mathbb{R}^{n_i}\times\mathbb{R}^{m_i}\to\mathbb{R}^{n_i}$ and $g_i:\mathbb{R}^{n_i}\times\mathbb{R}^{\bar{n}_i}\times\mathbb{R}^{m_i}\to\mathbb{R}^{n_i}$ represent the known local and unknown interconnection functions respectively, $\eta_i:\mathbb{R}^{n_i}\times\mathbb{R}^{m_i}\to\mathbb{R}^{n_i}$ denotes the system uncertainties, and $h_i:\mathbb{R}^{n_i}\times\mathbb{R}^{m_i}\to\mathbb{R}^{n_i}$ is the local fault function or fault dynamics.

The time profile $\Pi_i(k-k_0)$ is modeled by

$$\Pi_i(k-k_0)=diag\left\{\Omega_{i_1}(k-k_0),\Omega_{i_2}(k-k_0),\ldots,\Omega_{i_{n_i}}(k-k_0)\right\}$$

where

$$\Omega_{i_j}(\tau)=\begin{cases}0, & if\ \tau<0\\ 1-e^{-\bar{\kappa}_j\tau}, & if\ \tau\geq 0\end{cases}\quad for\ j=1,\ldots,n_i$$

is the fault profile and $\bar{\kappa}_j$ is an unknown constant that represents the rate at which a fault occurs. A larger value of $\bar{\kappa}_j$ indicates that it is an abrupt fault. The use of such time profiles is common in fault diagnosis literature [16,17]. Next standard assumptions are needed in order to proceed.

Assumption 1 [4]: The modeling uncertainty is bounded, i.e.$\|\eta_i(x_i(k),u_i(k))\|\leq\eta_{i_M},\ \forall(x,u)\in(\chi\times U),i=1,2,\ldots,N$, where η_{i_M} is a positive known constant.

Remark 1: Assumption 1 is needed to distinguish between faults and system uncertainties.

Assumption 2 [16]: Interconnection functions and fault functions are expressed as nonlinear in the unknown parameters (NLIP) [18], i.e. they can be approximated by two-layer NN with bounded activation functions and weight parameters.

Assumption 3 [9]: The interconnection terms are bounded by polynomial-type nonlinearities as $\|g_i(x(k), u_i(k))\| \leq \sum_{j=1}^{N}(\zeta_{ij}^0 + \zeta_{ij} x_j), i = 1,2, \dots, N$. For such system, it is considered that there exists a controller that is capable of keeping the system states bounded during healthy operating conditions. Thus, the bound on interconnection function given above in the healthy operating condition can be rewritten as $\|g_i(x(k), u_i(k))\| \leq g_{i_M}, i = 1,2, \dots, N$, where $\bar{g}_{i_M}$ is a positive constant.

Remark 2: The second part of the Assumption 3 is needed only in the healthy operating conditions, to analytically derive the fault detection threshold, whereas it is not used to prove the stability of the local FD observers. Boundedness of the interconnection term during healthy operating condition is mild since a number of decentralized control techniques [10,12] demonstrate stability of such systems. On the other hand, during fault conditions, the interconnection term being bounded above is no longer needed since it will not hold. Instead, the online approximator to be defined in the next section approximates the interconnection function as it propagates the fault from one subsystem to the other.

Assumption 4: It is assumed in this paper that faults occur one at a time in the subsystems. This assumption is required for partial isolation of faults.

3. FAULT DIAGNOSIS AND PROGNOSIS SCHEME

In the first part of this section, the proposed fault detection scheme is introduced.

3.1. FAULT DETECTION (FD)

In order to monitor the system states, local estimators which only use local measurements are designed. Since the interconnection functions and fault functions are not known, an online approximator is incorporated in each local estimator to approximate these functions. Unlike other fault detection schemes where the OLAD is turned on only after the detection, the online approximators used in our proposed estimators are always

turned on, in order to learn the possible fault dynamics as well as the interconnection dynamics.

Let ω_i be defined as the summation of interconnection term and fault function in subsystem i

$$\omega_i(k) = g_i\left(x_i(k), \bar{x}_i(k), u_i(k)\right) + \Pi\left(k - k_0\right) h_i\left(x_i(k), u_i(k)\right)$$

It is clear, based on (1), that the interconnection term at time k, will affect the local states at the next time instant k+1.Using this fact, the interconnection term at time k can be represented as a function of local states at time k+1 and local states and inputs at time k. Thus, $\omega_i(k)$ can be approximated by an online approximator such as a two layer neural network (NN) whose inputs consist of $x_i(k)$, $u_i(k)$, and $x_i(k+1)$, with bounded weights and approximation error, i.e.

$$\omega_i(k) = \theta_i^T\left(k\right)\phi_i\left(x_i\left(k+1\right), x_i\left(k\right), u_i\left(k\right)\right) + \varepsilon_i(k)$$

where $\theta_i(k)$ is the unknown parameter matrix, $\phi_i\left(x(k), u(k)\right)$ is a basis function like sigmoid, and $\varepsilon_i(k)$ is the approximation error which is bounded by ε_{i_M}. However, since the measured state vector, $x_i(k+1)$, is not available at time k, we will consider the online approximator one time step behind the actual system, in order to make the proposed scheme practical. Thus, the NN approximator will be incorporated in a nonlinear observer which is designed to work one time step behind the actual system. The residual, which is defined as the error between measured and estimated states, will then be used to update the NN weights.

Remark 3: The target weights are assumed to be time variant, because even when the interconnection term is time invariant, the occurrence and evolution of fault will make $\omega_i(.)$ a time variant function.

Consider the local nonlinear estimator for the ith subsystem described by

$$\begin{aligned}\hat{x}_i\left(k\right) = \lambda\hat{x}_i\left(k-1\right) + f_i\left(x_i\left(k-1\right), u_i\left(k-1\right)\right) - \lambda x_i\left(k-1\right) \\ + \hat{\omega}_i\left(x_i\left(k\right), x_i\left(k-1\right), u_i\left(k-1\right); \hat{\theta}_i\left(k-1\right)\right)\end{aligned} \quad (2)$$

for $k \geq 1$, where $\hat{x}_i(k) \in \mathbb{R}^{n_i}$ is the estimated local state vector of the ith subsystem, $\hat{\omega}_i: \mathbb{R}^{m_i} \times \mathbb{R}^{n_i} \times \mathbb{R}^{p_i \times n_i} \to \mathbb{R}^{n_i}$ is the output of the OLAD with $\hat{\theta}_i \in \mathbb{R}^{p_i \times n_i}$ being its set of unknown parameters and λ is a user defined constant, which must be selected in a way

that the eigenvalues of the closed loop system lie within the unit disc [18]. Initial values of the local fault detection (FD) estimator are taken to be $\hat{x}_i(0) = \hat{x}_{i_0}$, $\hat{\theta}_i(0) = \hat{\theta}_{i_0}$.

During the healthy operating conditions of the system, the following inequality holds $\|\omega_i(k)\| \le \|g_i(x(k), u_i(k))\|$ so that $\|\omega_i(k)\|$ remains bounded based on Assumptions 3. When a fault occurs, the magnitude of fault function and interconnection function will increase. Therefore, a fault can be detected by comparing the norm of OLAD output, $\|\hat{\omega}_i\|$, with a detection threshold ρ_i which will be defined later by using the bound on the interconnection functions in the healthy conditions as well as the bound on the OLAD approximation error. This is in contrast with detecting a fault by using the residual or state estimation error.

To move forward, define the i[th] subsystem residual as $e_i(k) = x_i(k) - \hat{x}_i(k)$. Prior to the occurrence of a fault, the local residual dynamics are given by

$$\begin{aligned} e_i(k) = \lambda e_i(k-1) + g_i(x(k-1), u_i(k-1)) + \eta_i(x_i(k-1), u_i(k-1)) \\ - \hat{\omega}_i(x_i(k), x_i(k-1), u_i(k-1); \hat{\theta}_i(k-1)) \end{aligned} \tag{3}$$

The next step in the design is to determine an update law for the online approximators which is given by

$$\begin{aligned} \hat{\theta}_i(k) = \hat{\theta}_i(k-1) + \alpha_i \phi_i(k-1) e_i^T(k-1) \\ - \gamma_i \| I - \alpha_i \phi_i(k-1) \phi_i^T(k-1) \| \hat{\theta}_i(k-1) \end{aligned} \tag{4}$$

where α_i>0 is the learning rate, $0 < \gamma_i < 1$ is the forgetting factor, and $\phi_i(k) = \phi_i(x_i(k), u_i(k))$ is a basis function such as sigmoid or RBF. Then, the output of the OLAD is calculated as

$$\hat{\omega}_i(k-1) = \hat{\theta}_i^T(k-1) \phi_i(x_i(k), x_i(k-1), u_i(k-1)) \tag{5}$$

Upon detection the local error dynamics can be derived by comparing (1) and (2) at time *k* as

$$\begin{aligned} e_i(k) = \lambda e_i(k-1) + g_i(x(k-1), u_i(k-1)) + \eta_i(x_i(k-1), u_i(k-1)) \\ + h_i(x_i(k-1), u_i(k-1)) - \hat{\omega}_i(x_i(k), x_i(k-1), u_i(k-1); \hat{\theta}_i(k-1)) \end{aligned} \tag{6}$$

Asserting the NLIP assumption on the local fault function, the above equation can be rewritten as

$$e_i(k) = \lambda e_i(k-1) + \eta_i\left(x_i(k-1), u_i(k-1)\right) + \tilde{\theta}_i^T(k-1)\phi_i\left(x_i(k), x_i(k-1), u_i(k-1)\right) + \varepsilon_i(k-1) \quad (7)$$

where $\tilde{\theta}_i(k) = \theta_i(k) - \hat{\theta}_i(k)$ represents the parameter estimation error and $\varepsilon_i(k)$ is the OLAD approximation error, which is bounded by ε_{i_M} as a result of assumption 2.

Next the stability of the local FD residual and parameter estimation errors is discussed.

Theorem 1 (Local Fault Detection Observer Performance): Let the proposed local FD observer defined in (2) be used to monitor the subsystem described by (1), and let the update law in (4) be used to update the unknown parameter vector, $\hat{\theta}_i(k)$. In the presence of system uncertainties, the local FD residual, $e_i(k)$, and the parameter estimation error, $\tilde{\theta}_i(k)$, are uniformly ultimately bounded, provided the user-defined constants, λ and α_i, and the basis function $\phi(.)$, are selected such that $|\lambda| < 0.5$, $\alpha_i < \sqrt{(1-4\lambda^2)/3\phi_{i_{max}}^2}$, and $\phi_{i_{max}} < 0.5$.

Proof: Refer to the appendix.

Theorem 1 guarantees the stability of the local FD residual and parameter estimation errors provided the design parameters are selected as derived in the theorem. As a result the interconnection functions can be approximated during the healthy conditions using only the local measurements in each subsystem. When a fault happens, the fault function is also approximated in the subsystem where it has occurred. Although the fault function only exists in one of the subsystems, it will affect the other subsystems through the interconnection term. Therefore, the estimation of interconnection functions in non-faulty subsystems, allows determination of non-local fault effects. This feature is also used for partial isolation of faults.

Theorem 2 (Robustness and Detectability): Consider the nonlinear subsystem defined by (1) and the local observer (2). No fault is detected under healthy operating conditions if the detection threshold is selected as

$$\rho_i = g_{i_M} + \varepsilon_{i_M} + \phi_{i_{\max}} \sqrt{D / C_2} + q_i \tag{8}$$

where q_i is a small positive constant. On the other hand, the fault in subsystem i will be detected by its local fault detector, if there exists a time instant k_d, at which the following condition on the fault function is satisfied

$$\left\| h_i(x_i(k_d), u_i(k_d)) \right\| \geq 2\rho_i - q_i \tag{9}$$

Proof: Refer to the appendix.

Remark 4: In the case of a nonlocal fault, the interconnection term causes the local residual to exceed its threshold by carrying the effects of nonlocal fault from a different subsystem. In this case, there is no local fault function and the online approximator will try to follow the change in the interconnection term caused by a nonlocal fault. Thus the local OLAD will not provide the approximation of a nonlocal fault function and thus TTF calculation is not accurate. Therefore, upon isolation of a nonlocal fault, the TTF module will not be continued.

Next section will discuss the partial fault isolation, upon detection of a fault by a local FD. With the proposed isolation method, the detected fault can be characterized as local or non-local fault.

3.2. FAULT ISOLATION

As discussed earlier, the online approximator in faulty subsystem estimates both the local fault function and the interconnection term, while the OLADs in other subsystems estimate only their interconnection function, which could be affected by the nonlocal fault. Therefore, not only the output of OLAD where the fault has occurred will increase above the detection threshold, but also the outputs of other OLADs can possibly increase due to interconnection effects. Thus, detection of fault could happen in more than one subsystem. However, it is important to note that local faults affect local measurements quicker than the non-local faults and they have shorter propagation delay. There is a stronger correlation between the local fault magnitude and corresponding

OLAD output magnitude. Using this fact, a fault can be isolated as local or nonlocal to a particular subsystem based on the time when it is detected.

In the proposed isolation method, communication between the LFD and the centralized isolation unit is required. However there is no need for the transmission of the measured or estimated state vector of all the subsystem. The only information which must be transmitted is the detection time in each local fault detector and there is no need for the detection information to be transmitted at each and every time instant. In fact this information must be sent from all the subsystems to a central isolation unit at time instants $k = jn$ where $j = 1,2, \ldots$ and n is a positive integer which determines the rate at which detection information must be collected from all the subsystems. In other words, the time interval between two consecutive transmissions will be equal to nT where T is the sampling time. Larger value of n will result in fewer number of transmissions over the network, while smaller value of n leads to faster isolation of faults. So there is a tradeoff here which means that n should be selected according to both the required isolation speed and preferred transmission interval in a specific system.

To formulate the isolation scheme, let $t_D^{(i)}$ be the variable used to store the detection information of subsystem *i* and let $t_D^{(i)}(0) = 0$. The value of $t_D^{(i)}$ will remain at zero unless a fault is detected by the LFD of subsystem *i*. Once a fault is detected by this LFD, $t_D^{(i)}$ will be set to the detection time, i.e., $t_D^{(i)}(k) = Tk_{detection}^{(i)}\ for\ k \geq k_{detection}^{(i)}$. Note that $t_D^{(i)}(k)$ is sent to the central isolation unit only when $k = nj$ where $j = 1,2, \ldots$.

The fault isolation flowchart is depicted in Figure 3.1. Once detection information $\left(t_D^{(i)}\right)$ is sent to the isolation unit by all the subsystems, the minimum among all of the nonzero detection times is calculated. Then for each subsystem, say subsystem *i*, $t_D^{(i)}$ is first compared to zero. When $t_D^{(i)}$ is equal to zero obviously no fault has been detected in subsystem *i*. However, when $t_D^{(i)} > 0$ a local or nonlocal fault has been detected in subsystem *i*. In this case, if the detection time $t_D^{(i)}$ is equal to the minimum of all nonzero detection times, then the fault will be isolated local to subsystem *i*, otherwise the fault is a nonlocal fault which has propagated to subsystem *i*.

Remark 5: Note that with this method of decentralized fault detection and isolation, not only the location of fault can be determined, but also all the subsystems which are affected by this fault are identified.

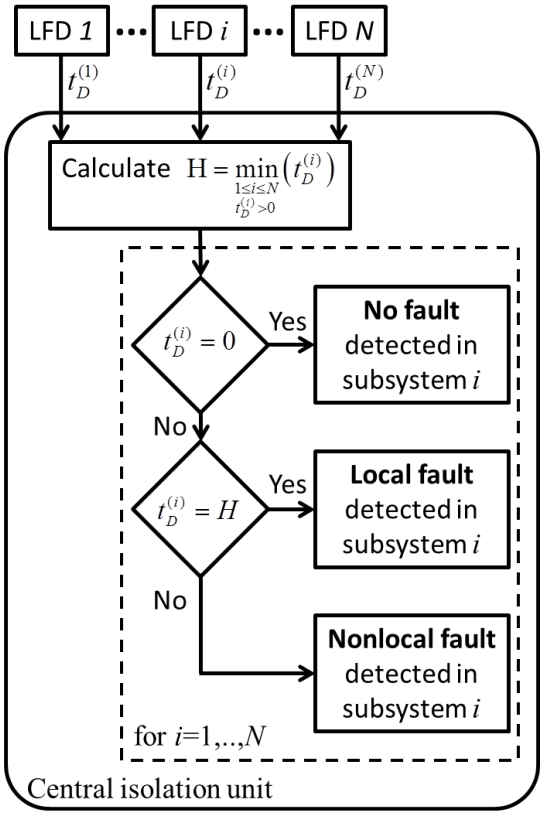

Figure 3.1. Flowchart of the fault isolation

3.3. FAILURE PREDICTION

The TTF determination is necessary for prognostics. This is also referred to as remaining useful life of the system. After the detection of a fault, by comparing the estimated states obtained from the observer to the user defined limits, time to failure can be determined [16]. The TTF is defined as the time elapsed when the first state reaches its limit. As mentioned before, a fault might be detected in more than one subsystem, since any local fault can influence other subsystems as well. Therefore, time-to-failure estimation should be performed for all the subsystems which are significantly affected by

the fault, i.e. all subsystems where detection has occurred. TTF estimation starts in a subsystem immediately after detection.

In order to predict the time of failure, the dynamics of the system can be used which will help determine the rate of change of system states. Since there exist unknown terms in the actual system dynamics (1), the observer dynamics (2) is utilized. According to the stability analysis presented earlier, the observer states follow the actual states with a bounded error which can be decreased by proper selection of design parameters.

Therefore, in the TTF determination, the estimated state dynamics in (2) are utilized to project the estimated state to reach a predefined threshold. The estimated state is driven by the fault approximator. The following theorem provides an analytical formula for finding time-to-failure.

Theorem 3 (TTF Estimation): Upon detection in subsystem *i*, TTF for the j^{th} state at the k^{th} time instant can be estimated using

$$TTF_{i,j} = \log_{\lambda}\left|\frac{(1-\lambda)x_{i,j_M} - v_{i,j}(k-1)}{\lambda(1-\lambda)\hat{x}_{i,j}(k-1) - \lambda v_{i,j}(k-1)}\right| \tag{10}$$

where x_{i,j_M} is the failure threshold of the jth state of the ith subsystem, $\hat{x}_{i,j}$ is the estimated value of the corresponding state, and $v_{i,j}(k-1)$ is the jth element of the vector $v_i(k-1)$ which is defined by

$$\begin{aligned} v_i(k-1) = {} & f_i\left(x_i(k-1), u_i(k-1)\right) - \lambda x_i(k-1) \\ & + \hat{\omega}_i\left(x_i(k), x_i(k-1), u_i(k-1); \hat{\theta}_i(k-1)\right) \end{aligned} \tag{11}$$

Proof: Refer to the appendix.

Figure 3.2 illustrates the process of finding the TTF after a fault is detected in subsystem *i*. At each time instant, after calculating the TTF for all of the local subsystem parameters, one should take the minimum of time to failure for all of the parameters, to get the overall TTF for the subsystem. This is because the system will be unsafe even if only one of its parameters reaches its limit.

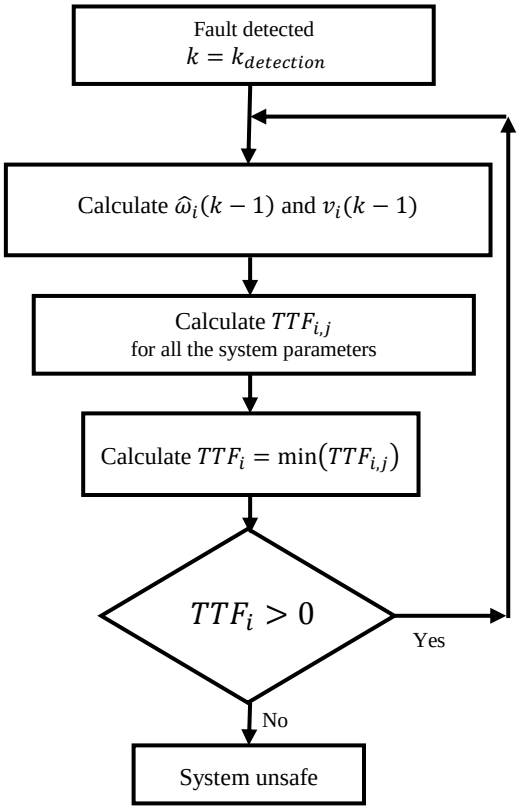

Figure 3.2. Flow chart of the TTF determination

4. SIMULATION RESULTS

A five-tank water system [13] which is shown in Figure 4.1, is considered to verify the proposed decentralized fault diagnosis scheme. This system has two input pumps with five connected water tanks, and can be decomposed into two subsystems; subsystem 1 includes tanks 1, 2, and 3, whereas subsystem 2 includes tanks 3, 4, and 5.

The system dynamics are described by [13]

$$x_1^{(1)}(k+1) = \frac{T}{A}\Big(u_1 - cs.sign\left(x_1^{(1)}(k) - x_1^{(2)}(k)\right) \cdot\sqrt{2g\left|x_1^{(1)}(k) - x_1^{(2)}(k)\right|}\Big) + \eta_1 + x_1^{(1)}(k)$$

$$x_1^{(2)}(k+1) = \frac{T}{A}\Big(cs.sign\left(x_1^{(1)}(k) - x_1^{(2)}(k)\right)\sqrt{2g\left|x_1^{(1)}(k) - x_1^{(2)}(k)\right|} - cs.sign\left(x_1^{(2)}(k) - x_1^{(3)}(k)\right).\sqrt{2g\left|x_1^{(2)}(k) - x_1^{(3)}(k)\right|}\Big) + \eta_2 + x_1^{(2)}(k)$$

$$x_1^{(3)}(k+1)=\frac{T}{A}\Big(cs.sign\Big(x_1^{(2)}(k)-x_1^{(3)}(k)\Big)\sqrt{2g\left|x_1^{(2)}(k)-x_1^{(3)}(k)\right|}$$
$$-cs.sign\Big(x_1^{(3)}(k)-x_2^{(2)}(k)\Big).\sqrt{2g\left|x_1^{(3)}(k)-x_2^{(2)}(k)\right|}\Big)+\eta_3+x_1^{(3)}(k)$$

$$x_2^{(2)}(k+1)=\frac{T}{A}\Big(cs.sign\Big(x_2^{(1)}(k)-x_2^{(2)}(k)\Big)\sqrt{2g\left|x_2^{(1)}(k)-x_2^{(2)}(k)\right|}$$
$$-cs.sign\Big(x_2^{(2)}(k)-x_2^{(3)}(k)\Big).\sqrt{2g\left|x_2^{(2)}(k)-x_2^{(3)}(k)\right|}\Big)+\eta_4+x_2^{(2)}(k)$$

$$x_2^{(3)}(k+1)=\frac{T}{A}\Big(u_2+cs.sign\Big(x_2^{(2)}(k)-x_2^{(3)}(k)\Big)$$
$$.\sqrt{2g\left|x_2^{(2)}(k)-x_2^{(3)}(k)\right|}-cs.\sqrt{2gx_2^{(3)}(k)}\Big)+\eta_5+x_2^{(3)}(k)$$

where $x_1(k)=\left[x_1^{(1)}(k),x_1^{(2)}(k),x_1^{(3)}(k)\right]^T$ is the first subsystem state vector, $x_2(k)=\left[x_2^{(1)}(k),x_2^{(2)}(k),x_2^{(3)}(k)\right]^T$ is the second subsystem state vector, T is the sampling time chosen to be 0.1 seconds, $A=0.0154\,m^2$ is the cross section of the tanks, $s=5\times10^{-5}\,m^2$ is the cross section of the connecting pipes, $c=1$ is the outflow coefficient, and $g=9.8\,m/s^2$ is the standard gravity. Note that the two subsystems share one of the states, i.e. $x_1^{(3)}=x_2^{(1)}$. Moreover, $\eta\big(x(k)\big)=[\eta_1(k)\ \eta_2(k)\ \eta_3(k)\ \eta_4(k)\ \eta_5(k)]^T$ represents the modeling uncertainty and is defined by

$$\eta=[10^{-2}\sin(0.7k)\ \ 10^{-2}\cos(0.8k)\ \ 10^{-2}\cos(0.5k)\ \ 10^{-2}\sin(0.6k)\ \ 10^{-2}cos(0.7k)]^T$$

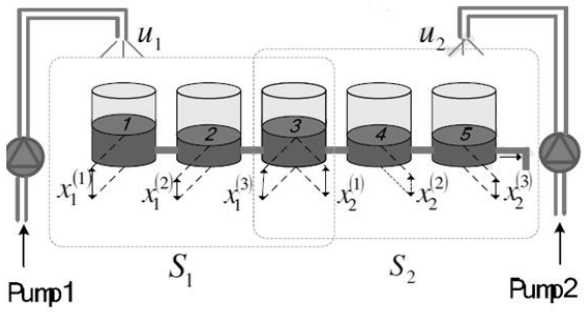

Figure 4.1. Five tank benchmarking system

An incipient actuator fault in pump 1 (located in subsystem 1) is seeded at time $t_0=50$s. The dynamics of actuator fault in subsystem 1 is described by

$$\Pi(k-k_0)h_1\big(x_1(k),u_1(k)\big)=\left[\left(1-e^{-(k-k_0)}\right)\left(\frac{sgn[k-k_0]+1}{2}\right)\frac{9T}{A}u_1(k),\ 0,\ 0\right]^T$$

Online approximators $\widehat{\omega}_1$ and $\widehat{\omega}_2$ are both made up of 7-input 3-output neural networks which consists of 8 basis functions. The basis functions are sigmoid type and they satisfy $\phi_{i_{max}} < 0.5$. The inputs are local states at current and next time instants and the local input. The estimator and adaptive law parameters are taken as $\alpha_i = 0.1, \gamma_i = 10^{-5}$, $and\ \lambda = 0.01$. The bounds on the uncertainty and interconnection terms are $\eta_{i_M} = 0.029$ and $g_{i_M} = 0.022$. In order to calculate the detection thresholds from (8), the maximum neural network approximation errors ε_{i_M} are required. Unless the interconnection term is represented as a linear function of states and inputs, ε_{i_M} cannot be found analytically. However, we know that the approximation error in healthy conditions is definitely less than the upper bound on interconnection term. Thus, we will replace ε_{i_M} by g_{i_M}. By using these parameter values in (8), the detection thresholds are calculated as $\rho_i = 0.09$.

Figures 4.2 and 4.3 show the actual states of subsystem 1 and 2, respectively. Obviously, the behavior of the system changes due to fault which starts to evolve 50 seconds into the simulations. As mentioned previously, the online approximators are always online to learn the interconnection dynamics in all subsystems. After the occurrence of fault, the OLAD in faulty subsystem will also approximate the fault dynamics. Norms of the outputs of both OLADs are plotted along with the detection threshold in Figure 4.4. As expected, the fault in subsystem 1 not only affects the local states, but also affects the interconnection terms. That is why the output of both OLADs increase after occurrence of fault. However, the growth rate of the output of OLAD in subsystem 1 (where the fault is initiated) is significantly higher than the growth rate of the output of OLAD in subsystem 2. Thus the fault is detected first in subsystem 1.

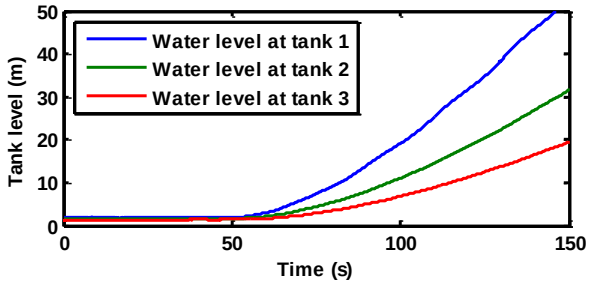

Figure 4.2. States of subsystem 1

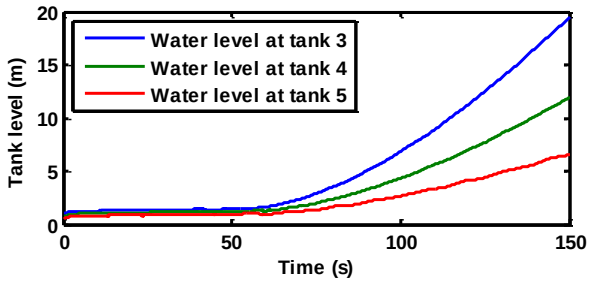

Figure 4.3. States of subsystem 2

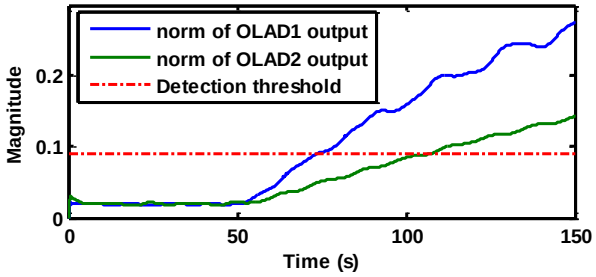

Figure 4.4. OLAD outputs and detection threshold

Local residuals, which are generated by comparing the actual and estimated subsystem states, are mainly used for updating the NN weights. The norm of local residual of both subsystems is plotted in Figure 4.5. Norm of residuals is small and bounded both before and after the fault, which shows the boundedness of the state estimation errors due to successful estimation of unknown dynamics by the stable weight update laws.

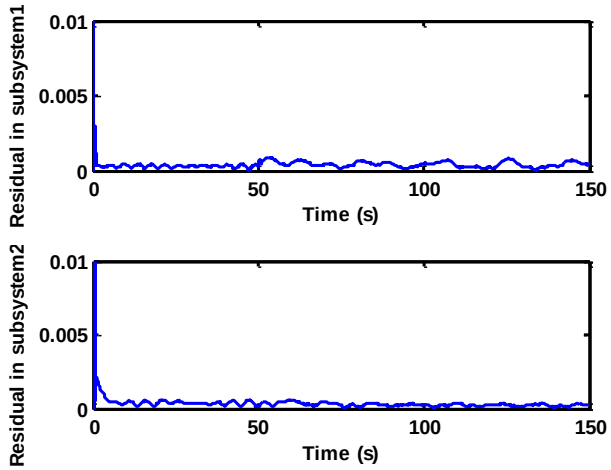

Figure 4.5. Residuals in subsystem 1 and subsystem 2

The output of OLAD in subsystem 1, $\widehat{\omega}_1(.)$, is plotted along with the actual values of the function $\omega_1(.)$ over time in Figure 4.6. The OLAD reasonably tracks the unknown vector function $\omega_1(.)$ and this allows accurate estimation of time-to-failure which is done based on the estimated parameters and their failure thresholds.

Figures 4.7 and 4.8 show the estimated time-to-failure for subsystem 1 and subsystem 2 respectively when the failure simulation is accelerated. Time-to-failure is calculated for each state based on the proposed algorithm, and then the subsystem time-to-failure is obtained by taking the minimum among estimated TTF for all states of the corresponding subsystem. The TTF of subsystem 1 approaches zero faster than subsystem 2, because the fault is seeded in subsystem 1 and it has an attenuated and delayed effect on second subsystem. It should be noted that, the whole system should be stopped before the TTF in any subsystem reaches zero. In this example, the operation of system is unsafe after t=100.2s where TTF for subsystem 1 reaches zero.

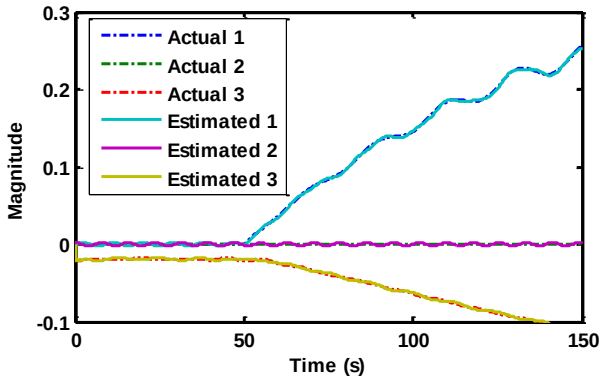

Figure 4.6. Actual and estimated magnitude of $\omega_1(k)$

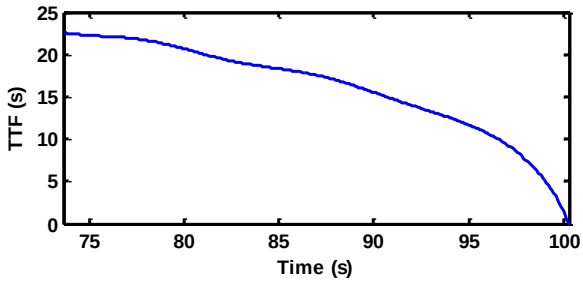

Figure 4.7. TTF vs. Time for first subsystem 1

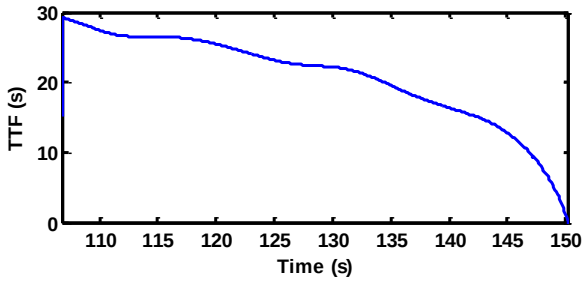

Figure 4.8. TTF vs. Time for first subsystem 2

5. CONCLUSIONS

The proposed decentralized fault prognosis scheme renders satisfactory performance when the faults are either local or non-local to the subsystem. Only the local subsystem states were used and partial isolation of faults is possible. Fault can be detected in all the subsystems that are significantly affected. Consequently, local and non-local faults can be distinguished by using a central isolation unit. Upon detection in each subsystem, time to failure can be predicted by using the estimated state dynamics driven by the fault approximation. No prior offline training or fault data is necessary in order to detect or isolate faults whereas local state measurements are utilized for detection and isolation. Hence, this scheme can save both time and cost while it is easily implementable on embedded system by measuring the local subsystem states. The simulation results verify the effectiveness of this approach.

6. REFERENCES

[1] S. Dash and V. Venkatasubramanian, "Challenges in the industrial applications of fault diagnostic systems," *Comput. Chem. Eng*., vol. 24, no. 2–7, pp. 785–791, 2000.

[2] R. Isermann, "Model-based fault-detection and diagnosis–status and applications," *Annual Reviews in Control*, vol. 29, pp. 71-85, 2005.

[3] H. Wang and S. Daley, "Actuator fault diagnosis: an adaptive observer-based technique," *IEEE Transactions on Automatic Control*, vol.41, no.7, pp. 1073-1078, 1996.

[4] M. A. Demetriou and M. M. Polycarpou, "Incipient fault diagnosis of dynamical systems using online approximators," *IEEE Trans. on Automatic Control,* vol. 43, no. 11, pp. 1612-1617, 1998.

[5] H. Ferdowsi and S. Jagannathan, "A unified model-based fault diagnosis scheme for nonlinear discrete-time systems with additive and multiplicative faults," *50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC),* pp. 1570-1575, 2011.

[6] Y. Maki and K.A. Loparo, "A neural-network approach to fault detection and diagnosis in industrial processes," *IEEE Transactions on Control Systems Technology,* vol. 5, no. 6, pp. 529-541, 1997.

[7] A. Bernieri, M. D'Apuzzo, L. Sansone, and M. Savastano, "A neural network approach for identification and fault diagnosis on dynamic systems," *IEEE Transactions on Instrumentation and Measurement,* vol. 43, no. 6, pp. 867-873, 1994.

[8] D. Blake and M. Brown, "Simultaneous, multiplicative actuator and sensor fault estimation," *Fuzzy Systems Conference 2007, FUZZ-IEEE,* pp. 1-6, 2007.

[9] C. J. Lopez-Toribio and R. J. Patton, "Fuzzy observers for nonlinear dynamic systems fault diagnosis," *Proc. IEEE Conf. Decision Control*, pp. 84–89, 1998.

[10] S. N. Huang, K. K. Tan, and T. H. Lee, "Decentralized control of a class of large-scale nonlinear systems using neural networks," *Automatica,* vol. 41, pp. 1645-1649, 2005.

[11] J. D. Boskovic and R. K. Mehra, "A decentralized fault-tolerant control system for accommodation of failures in higher-order flight control actuators," *IEEE Trans on Control Systems Technology,* Vol. 18, no. 5, pp. 1103-1115, 2010.

[12] S. N. Huang, K. K. Tan, and T. H. Lee, "Nonlinear adaptive control of interconnected systems using neural networks," *IEEE Trans on Neural Networks*, vol. 17, no. 1, pp. 243-246, 2006.

[13] R. M. G. Ferrari, T. Parisini, and M. M. Polycarpou, "Distributed Fault Diagnosis with Overlapping Decompositions: An Adaptive Approximation Approach," *IEEE Trans on Automatic Control*, vol. 54, no. 4, pp. 794–799, 2009.

[14] S. Stankovic, N. Ilic, Z. Djurovic, M. Stankovic, K.H. Johansson, "Consensus based overlapping decentralized fault detection and isolation," *Proc. 2010 Conference on Control and Fault-Tolerant Systems (SysTol'10),* pp. 570–575, 2010.

[15] F. Caccavale and L. Villani, "An adaptive observer for fault diagnosis in nonlinear discrete-time systems," *Proc. of the American Control Conference,* pp. 2463-2468, 2004.

[16] B. T. Thumati and S. Jagannathan, "A model based fault detection and prediction scheme for nonlinear multivariable discrete-time systems with asymptotic stability guarantees," *IEEE Transactions on Neural Networks*, vol. 21, no. 3, pp. 404-423, 2010.

[17] J. Zhang and A. J. Morris, "On-line process fault diagnosis using fuzzy neural networks," *Intelligent Systems Engineering*, pp. 37-47, 1994.

[18] S. Jagannathan, *Neural Network Control of Nonlinear Discrete –time Systems*, CRC publications, NY, 2006.

[19] G. H. Golub and C. F. V. Loan, *Matrix computations (3rd ed.)*, Johns Hopkins University Press, 1996.

7. APPENDIX

Proof of Theorem 1: Consider the following Lyapunov function candidate

$$V = e_i^T(k-1)e_i(k-1) + tr\left\{\tilde{\theta}_i^T(k-1)\tilde{\theta}_i(k-1)\right\} \tag{A.1}$$

The Lyapunov function is deliberately selected at time $k-1$, because the observer is one time step behind the actual system and its output is not available at time k. In other words, $k-1$ is the current time instant for the observer. The first difference of the Lyapunov function is given by

$$\Delta V = \underbrace{\left(e_i^T(k)e_i(k) - e_i^T(k-1)e_i(k-1)\right)}_{\Delta V_1} + \underbrace{tr\left\{\tilde{\theta}_i^T(k)\tilde{\theta}_i(k) - \tilde{\theta}_i^T(k-1)\tilde{\theta}_i(k-1)\right\}}_{\Delta V_2} \tag{A.2}$$

Substitute $e_i(k)$ from the local error dynamics (7), in ΔV_1 to get

$$\begin{aligned}\Delta V_1 = \Big[& \lambda e_i(k-1) + \eta_i\left(x_i(k-1), u_i(k-1)\right) + \varepsilon_i(k-1) \\ & + \tilde{\theta}_i^T(k-1)\phi_i\left(x_i(k), x_i(k-1), u_i(k-1)\right)\Big]^T \\ .\Big[& \lambda e_i(k-1) + \eta_i\left(x_i(k-1), u_i(k-1)\right) + \varepsilon_i(k-1) \\ & + \tilde{\theta}_i^T(k-1)\phi_i\left(x_i(k), x_i(k-1), u_i(k-1)\right)\Big] - e_i^T(k-1)e_i(k-1)\end{aligned}$$

Use the Cauchy-Schwarz inequality ($(s_1+s_2+\ldots+s_n)^T(s_1+s_2+\ldots+s_n) \leq n(s_1^T s_1+s_2^T s_2+\ldots+s_n^T s_n)$) to arrive at

$$\begin{aligned}\Delta V_1 \le\ & 4\lambda^2 e_i^T(k-1)e_i(k-1) + 4\varepsilon_i^T(k-1)\varepsilon_i(k-1) \\ & + 4\eta_i^T\left(x_i(k-1),u_i(k-1)\right)\eta_i\left(x_i(k-1),u_i(k-1)\right) \\ & + 4\phi_i^T\left(k-1\right)\tilde{\theta}_i(k-1)\tilde{\theta}_i^T(k-1)\phi_i\left(k-1\right) - e_i^T(k-1)e_i(k-1) \qquad (A.3)\end{aligned}$$

Now substitute $\hat{\theta}_i(k)$ from (4), in ΔV_2

$$\begin{aligned}\Delta V_2 = tr\Big\{ & \Big[\left(1-\gamma_i\left\|I-\alpha_i\phi_i(k-1)\phi_i^T(k-1)\right\|\right)\tilde{\theta}_i(k-1) \\ & + \gamma_i\left\|I-\alpha_i\phi_i(k-1)\phi_i^T(k-1)\right\|\theta_i - \alpha_i\phi_i(k-1)e_i^T(k-1)\Big]^T \\ & .\Big[\left(1-\gamma_i\left\|I-\alpha_i\phi_i(k-1)\phi_i^T(k-1)\right\|\right)\tilde{\theta}_i(k-1) \\ & + \gamma_i\left\|I-\alpha_i\phi_i(k-1)\phi_i^T(k-1)\right\|\theta_i - \alpha_i\phi_i(k-1)e_i^T(k-1)\Big] - \tilde{\theta}_i^T(k-1)\tilde{\theta}_i(k-1)\end{aligned}$$

Applying the Cauchy-Schwarz inequality yields

$$\begin{aligned}\Delta V_2 \le\ & 3tr\Big\{\alpha_i^2 e_i(k-1)\phi_i^T(k-1)\phi_i(k-1)e_i^T(k-1) \\ & + \gamma_i^2\left\|I-\alpha_i\phi_i(k-1)\phi_i^T(k-1)\right\|^2\theta_i^T(k-1)\theta_i(k-1) \\ & + \left(1-\gamma_i\left\|I-\alpha_i\phi_i(k-1)\phi_i^T(k-1)\right\|\right)^2\tilde{\theta}_i^T(k-1)\tilde{\theta}_i(k-1)\Big\} - \tilde{\theta}_i^T(k-1)\tilde{\theta}_i(k-1) \\ \le\ & tr\Big\{2\tilde{\theta}_i^T(k-1)\tilde{\theta}_i(k-1) - 6\gamma_i\left\|I-\alpha_i\phi_i(k-1)\phi_i^T(k-1)\right\|\tilde{\theta}_i^T(k-1)\tilde{\theta}_i(k-1) \\ & + 3\gamma_i^2\left\|I-\alpha_i\phi_i(k-1)\phi_i^T(k-1)\right\|^2\tilde{\theta}_i^T(k-1)\tilde{\theta}_i(k-1)\Big\} \\ & + tr\Big\{3\gamma_i^2\left\|I-\alpha_i\phi_i(k-1)\phi_i^T(k-1)\right\|^2\theta_i^T(k-1)\theta_i(k-1)\Big\} \\ & + 3\alpha_i^2 e_i^T(k-1)e_i(k-1)\phi_i^T(k-1)\phi_i(k-1) \qquad (A.4)\end{aligned}$$

By combining ΔV_1 and ΔV_2 from (A.3) and (A.4), the difference of Lyapunov function can be represented by

$$\begin{aligned}\Delta V =\ & \Delta V_1 + \Delta V_2 \\ \le\ & 4\lambda^2 e_i^T(k-1)e_i(k-1) + 4\varepsilon_i^T(k-1)\varepsilon_i(k-1) \\ & + 4\eta_i^T\left(x_i(k-1),u_i(k-1)\right)\eta_i\left(x_i(k-1),u_i(k-1)\right) \\ & + 4\phi_i^T\left(k-1\right)\tilde{\theta}_i(k-1)\tilde{\theta}_i^T(k-1)\phi_i\left(k-1\right) - e_i^T(k-1)e_i(k-1) \\ & + tr\Big\{2\tilde{\theta}_i^T(k-1)\tilde{\theta}_i(k-1) - 6\gamma_i\left\|I-\alpha_i\phi_i(k-1)\phi_i^T(k-1)\right\|\tilde{\theta}_i^T(k-1)\tilde{\theta}_i(k-1) \\ & \quad + 3\gamma_i^2\left\|I-\alpha_i\phi_i(k-1)\phi_i^T(k-1)\right\|^2\tilde{\theta}_i^T(k-1)\tilde{\theta}_i(k-1)\Big\} \\ & + tr\Big\{3\gamma_i^2\left\|I-\alpha_i\phi_i(k-1)\phi_i^T(k-1)\right\|^2\theta_i^T(k-1)\theta_i(k-1)\Big\} \\ & + 3\alpha_i^2 e_i^T(k-1)e_i(k-1)\phi_i^T(k-1)\phi_i(k-1)\end{aligned}$$

Taking the Frobenius norm, and using the result of assumptions 1 and 2, it can be shown that

$$\begin{aligned}\Delta V \le & -(1-4\lambda^2)\left\|e_i(k-1)\right\|^2 + 3\alpha_i^2\phi_{i_{max}}^2\left\|e_i(k-1)\right\|^2 \\ & -\left(6\gamma_i\left\|I-\alpha_i\phi_i(k-1)\phi_i^T(k-1)\right\| - 2 - 4\phi_{i_{max}}^2 - 3\gamma_i^2\left\|I-\alpha_i\phi_i(k-1)\phi_i^T(k-1)\right\|^2\right)\tilde{\theta}_i(k-1)^2 \\ & + 4\eta_{i_M}^2 + 4\varepsilon_{i_M}^2 + 3\gamma_i^2\left\|I-\alpha_i\phi_i(k-1)\phi_i^T(k-1)\right\|^2\left\|\theta_i(k-1)\right\|^2\end{aligned}$$

Therefore,

$$\begin{aligned}\Delta V \le & -(1-4\lambda^2-3\alpha_i^2\phi_{i_{max}}^2)\left\|e_i(k-1)\right\|^2 \\ & -\left(6\gamma_i\left\|I-\alpha_i\phi_i(k-1)\phi_i^T(k-1)\right\| - 2 - 4\phi_{i_{max}}^2 - 3\gamma_i^2\left\|I-\alpha_i\phi_i(k-1)\phi_i^T(k-1)\right\|^2\right)\left\|\tilde{\theta}_i(k-1)\right\|^2 \\ & + 4\eta_{i_M}^2 + 4\varepsilon_{i_M}^2 + 3\gamma_i^2\left\|I-\alpha_i\phi_i(k-1)\phi_i^T(k-1)\right\|^2\theta_{i_{\max}}^2 \qquad (A.5)\end{aligned}$$

If the following conditions are satisfied (which are guaranteed with the selection of user-defined parameters as stated in the theorem)

$$1-4\lambda^2-3\alpha_i^2\phi_{i_{max}}^2 > 0$$

$$6\gamma_i\left\|I-\alpha_i\phi_i(k)\phi_i^T(k)\right\| - 3\gamma_i^2\left\|I-\alpha_i\phi_i(k)\phi_i^T(k)\right\|^2 - 2 - 4\phi_{i_{max}}^2 > 0$$

then $\Delta V \le 0$, if also

$$\left\|e_i(k-1)\right\| > \sqrt{D/C_1} \quad or \quad \left\|\tilde{\theta}_i(k-1)\right\| > \sqrt{D/C_2} \qquad (A.6)$$

where

$$\begin{aligned}C_1 &= 1-4\lambda^2-3\alpha_i^2\phi_{i_{max}}^2 \\ C_2 &= 6\gamma_i\left\|I-\alpha_i\phi_i(k-1)\phi_i^T(k-1)\right\| - 2 - 4\phi_{i_{max}}^2 - 3\gamma_i^2\left\|I-\alpha_i\phi_i(k-1)\phi_i^T(k-1)\right\|^2 \\ D &= 4\eta_{i_M}^2 + 4\varepsilon_{i_M}^2 + 3\gamma_i^2\left\|I-\alpha_i\phi_i(k-1)\phi_i^T(k-1)\right\|^2\theta_{i_{\max}}^2\end{aligned}$$

Therefore, with the appropriate choice of design parameters, the local FD residual, $e_i(k)$, and the parameter estimation error $\tilde{\theta}_i(k)$, will be uniformly ultimately bounded with the bounds given in (A.6).

Proof of Theorem 2: Consider the output of local OLAD in subsystem *i*

$$\begin{aligned}\hat{\omega}_i(k) &= \hat{\theta}_i^T(k)\phi_i\left(x_i(k+1), x_i(k), u_i(k)\right) \\ &= \theta_i^T(k)\phi_i\left(x_i(k+1), x_i(k), u_i(k)\right) - \tilde{\theta}_i^T(k)\phi_i\left(x_i(k+1), x_i(k), u_i(k)\right) \\ &= \omega_i(k) - \varepsilon_i(k) - \tilde{\theta}_i^T(k)\phi_i\left(x_i(k+1), x_i(k), u_i(k)\right)\end{aligned}$$

By taking Frobenius norm and using Assumptions 1 and 3 as well as the result of theorem 1, the limit on ω_i in healthy conditions is obtained as

$$\begin{aligned}\left\|\hat{\omega}_i(k)\right\| &\le \left\|g_i(x(k), u_i(k))\right\| + \left\|\varepsilon_i(k)\right\| + \phi_{i_{\max}}\left\|\tilde{\theta}(k)\right\| \\ &\le g_{i_M} + \varepsilon_{i_M} + \phi_{i_{\max}}\sqrt{D/C_2}\end{aligned} \qquad (A.7)$$

Therefore, if the detection threshold is selected as in (8), then no fault is detected as long as the system is working under healthy operating conditions.

To find the detectability condition, the output of OLAD in the faulty subsystem is utilized

$$\begin{aligned}\left\|\hat{\omega}_i(k_d)\right\| &\ge \left\|h_i(x_i(k_d), u_i(k_d))\right\| - \left\|g_i(x(k_d), u_i(k_d))\right\| - \left\|\varepsilon_i(k_d)\right\| - \phi_{i_{\max}}\left\|\tilde{\theta}(k_d)\right\| \\ &\ge \left\|h_i(x_i(k_d), u_i(k_d))\right\| - g_{i_M} - \varepsilon_{i_M} - \phi_{i_{\max}}\sqrt{D/C_2}\end{aligned}$$

Therefore, if there exist a time instant k_d at which the following condition is satisfied

$$\left\|h_i(x_i(k_d), u_i(k_d))\right\| - g_{i_M} - \varepsilon_{i_M} - \phi_{i_{\max}}\sqrt{D/C_2} \ge \rho_i \qquad (A.8)$$

or equivalently

$$\left\|h_i(x_i(k_d), u_i(k_d))\right\| \ge 2\rho_i - q_i$$

then the fault will be detected in the faulty subsystem.

Proof of Theorem 3: Consider the observer dynamics in (2) represented as

$$\hat{x}_i(k) = \lambda\hat{x}_i(k-1) + bv_i(k-1) \qquad (A.9)$$

where *b=1* and v_i , defined in (11), acts as the input to the linear system of (A.9). By assuming that the fault is detected at time k_d, the response to this set of linear state space equations at time $k_f > k_d$ is given by

$$\hat{x}_i(k_f) = \lambda^{k_f-k+1}\hat{x}_i(k-1) + \sum_{l=k-1}^{k_f-1}\lambda^{k_f-l-1}bv_i(l) \qquad (A.10)$$

With the assumption that $v_i(l) = v_i(k-1)$ for $k-1 \le l \le k_f$ (which is reasonable, since the fault is assumed to be incipient type), (A.10) can be rewritten as

$$\hat{x}_i\left(k_f\right) = \lambda^{k_f-k+1}\hat{x}_i\left(k-1\right) + bv_i\left(l\right)\sum_{l=k-1}^{k_f-1}\lambda^{k_f-l-1}$$

$$= \lambda^{k_f-k+1}\hat{x}_i\left(k-1\right) + bv_i\left(k-1\right)\frac{1-\lambda^{k_f-k+1}}{1-\lambda}$$

Now suppose that $k_{f_{i,j}}$ is the time when the j[th] state of subsystem *i*, reaches its failure threshold, i.e. $\hat{x}_{i,j}\left(k_{f_{i,j}}\right) = \bar{x}_{i,j_M}$.

$$x_{i,j_M} = \lambda^{k_{f_{i,j}}-k+1}\hat{x}_{i,j}\left(k-1\right) + v_{i,j}\left(k-1\right)\frac{1-\lambda^{k_{f_{i,j}}-k+1}}{1-\lambda}$$

$$\Rightarrow (1-\lambda)x_{i,j_M} - v_{i,j}\left(k-1\right)$$

$$= \lambda\lambda^{k_{f_{i,j}}-k}\left(\left(1-\lambda\right)\hat{x}_{i,j}\left(k-1\right) - v_{i,j}\left(k-1\right)\right)$$

$$\Rightarrow \frac{(1-\lambda)x_{i,j_M} - v_{i,j}\left(k-1\right)}{\lambda\left(1-\lambda\right)\hat{x}_{i,j}\left(k-1\right) - \lambda v_{i,j}\left(k-1\right)} = \lambda^{k_{f_{i,j}}-k}$$

Therefore, the time to failure for the j[th] state of the i[th] subsystem can be estimated by

$$TTF_{i,j} = k_{f_{i,j}} - k = \log_\lambda\left|\frac{(1-\lambda)x_{i,j_M} - v_{i,j}\left(k-1\right)}{\lambda\left(1-\lambda\right)\hat{x}_{i,j}\left(k-1\right) - \lambda v_{i,j}\left(k-1\right)}\right|$$

IV. A Decentralized Fault Detection and Accommodation Scheme for Interconnected Nonlinear Continuous-time Systems

Hasan Ferdowsi and S. Jagannathan

Abstract

In this paper, a novel decentralized fault detection and accommodation (FDA) methodology is proposed for interconnected nonlinear continuous-time systems by using local subsystem states alone in contrast with traditional distributed FDA schemes where the measured or the estimated state vector of the overall system is needed. The proposed decentralized FDA scheme uses local state and input vectors and minimizes the fault effects on all the subsystems. For this purpose, a network of local fault detectors (LFD) is proposed for fault detection where a fault is detected by generating a residual from the measured and estimated state vectors while the fault dynamics are estimated by using an online approximator (OLA) upon detection. Subsequently, a fault accommodation scheme is initiated in the subsystem by using a second OLA to augment the control input of each subsystem in order to minimize the effects of the faults on the overall system. Stability of both detection and accommodation schemes are discussed in the paper. Moreover time-to-accommodation is introduced. Finally the proposed methods are verified in the simulation environment.

1. INTRODUCTION

Many industrial systems such as power or water distribution networks, telecommunication networks, and so on are complex, large-scale, spatially distributed and interconnected nonlinear systems. Suitable fault detection schemes are required for these systems to ensure that a fault at any given location can be detected at an incipient stage in order to prevent catastrophic failure of the overall system. Although certain faults are critical and force the overall system to shut down, other faults at an incipient stage can be accommodated for a limited time.

Basically two types of fault detection schemes are available in the literature [1]: data-driven and model-based schemes. Model-based schemes are more desirable as they function online and in the absence of significant faulty or healthy data since they use a state estimator as a reference model for fault detection. State estimators are designed by using high gain observers [2], neural network observers [3], and geometric techniques [4]. On the other hand, a fault accommodation scheme is also available by utilizing observer-based schemes [5], adaptive estimators [6], and so on. However, these schemes are all centralized where the entire state vector is needed.

Recently design of distributed fault detection schemes for interconnected systems have been proposed by using H_∞ based method [7], multiblock kernel partial least squares [8], fuzzy observers [9] and so on. In addition, several distributed accommodation schemes [10-12] have been introduced for such interconnected systems. However, most of them [7-12] require either entire state or estimated state vector for each local fault detector (LFD) since these are merely distributed schemes. Typically, it is not always possible to provide the information of the entire state or its estimated value for LFDs of a large-scale spatially distributed system unless some information from other subsystems is communicated. Even if this is possible, the information will be delayed and outdated, besides being expensive.

In contrast to the aforementioned schemes [7-12], in this paper our objective is to design a LFD for each subsystem. Each LFD uses local states of that subsystem alone to detect a fault and approximate its dynamics. Upon detection, a fault accommodation scheme is proposed in which the control input of each subsystem is augmented in order to accommodate the effects of the fault by using the local subsystem states alone.

Each LFD is a nonlinear observer that acts as a reference model and it estimates the local subsystem states by using the known part of the system dynamics and inputs. A local residual signal is generated at each subsystem by comparing the estimated states with that of the actual subsystem states. Whenever the local detection residual exceeds its threshold, a fault is detected in that subsystem. Consequently an online approximator (OLA) or a fault detection OLA will be turned on in the LFD to approximate the local fault dynamics and the other LFDs are notified in order to keep all the OLAs in those subsystems offline, even if their local residual exceed the detection threshold afterwards. This is required because a fault in one subsystem can also affect other subsystems through the interconnections, but the actual fault that needs to be approximated only exists in the subsystem where the detection is performed first.

Upon detection, a novel fault accommodation scheme is initiated at each subsystem which will modify the control input vector of all subsystems to minimize the effect of the fault and force all system states to continue tracking their desired trajectories. A second OLA is also initiated in the faulty subsystem as well as all other subsystems for accommodation. The fault detection OLA output will also be utilized in the faulty subsystem to mitigate the local fault dynamics.

Since a fault in one subsystem can propagate to other subsystems via interconnection terms, an accommodation OLA will be initiated in each of the other subsystems to mitigate this effect. Lyapunov proofs are offered for the local fault detection and accommodation schemes. Furthermore, utilization of two OLAs for detection and accommodation in the faulty subsystem enables the proposed scheme to provide an estimation of fault dynamics and perform accommodation simultaneously. The estimation of fault dynamics can be further used for classification of fault type and time-to-failure (TTF) determination.

The TTF which is the time left until the system reaches its failure limit, is approximated by using the fault parameter update law. In addition, time to accommodation (TTA) is introduced. Estimation of TTF and TTA determines whether or not the accommodation unit can bring the system performance back to normal before the system reaches a failure. Simulation results verify theoretical claims.

The contributions of this paper include the development of a new decentralized fault detection and accommodation scheme for nonlinear interconnected continuous-time systems. Both the detection and accommodation schemes use local subsystem states and inputs to detect and accommodate the faults. Lyapunov stability analysis is included for these schemes. In addition, analytical formula for calculating TTF and TTA online is derived.

Next, the interconnected nonlinear continuous-time system description is presented followed by decentralized detection and accommodation schemes.

2. SYSTEM DESCRIPTION

Consider a nonlinear continuous-time system that is comprised of N interconnected subsystems. The dynamics of the i[th] subsystem with n_i states are described by

$$\begin{aligned} \dot{x}_{ij}(t) &= x_{i(j+1)}(t) \qquad\qquad j = 1, 2, \ldots, n_i - 1 \\ \dot{x}_{in_i}(t) &= f_i\left(x_i(t), u_i(t)\right) + g_i\left(x_i(t), \bar{x}_i(t)\right) + \eta_i\left(x(t)\right) + \Omega_i\left(t - t_0\right) h_i\left(x_i(t)\right) \end{aligned} \tag{1}$$

where $u_i \in \mathbb{R}$ is the local control input, $x_i \in \mathbb{R}^{n_i}$ is the local state vector, $\bar{x}_i \in \mathbb{R}^{\bar{n}_i}$ is the vector of interconnection states, $f_i : \mathbb{R}^{n_i} \times \mathbb{R} \to \mathbb{R}^{n_i}$ and $g_i : \mathbb{R}^{n_i} \times \mathbb{R}^{\bar{n}_i} \to \mathbb{R}$ represent the known local and unknown interconnection functions respectively, $\eta_i : \mathbb{R}^{n_i} \to \mathbb{R}$ denote the system uncertainties, and $h_i : \mathbb{R}^{n_i} \to \mathbb{R}$ is the local fault function.

The time profile of a fault $\Omega_i(t - t_0)$ is modeled by

$$\Omega_i\left(t - t_0\right) = \begin{cases} 0, & \text{if } t < t_0 \\ 1 - e^{-\bar{\kappa}_i (t - t_0)}, & \text{if } t \geq t_0 \end{cases}$$

where $\bar{\kappa}_i$ is an unknown constant that represents the rate at which a fault occurs. A larger value of $\bar{\kappa}_i$ indicates that the fault is an abrupt fault while small values of $\bar{\kappa}_i$ indicate that the fault is of an incipient type. The use of such time profiles is common in the fault diagnosis literature [5]. Next three standard assumptions from the literature are presented.

Assumption 1: The modeling uncertainty is bounded, i.e. $\left\|\eta_i\left(x(t)\right)\right\| \leq \eta_{i_M}$, $\forall (x, u) \in (\chi \times U), i = 1,2, \ldots, N$, where η_{i_M} is a known positive constant.

Remark 1: Assumption 1 is required to differentiate between faults and modeling uncertainties and to analytically define fault detection thresholds to prevent false or missed alarms.

Assumption 2: A fault occurs only in one of the subsystems and the fault functions can be expressed as nonlinear in the unknown parameters (NLIP) [13]. The NLIP representation for fault functions allows the use of two-layer NNs with nonlinear activation functions.

Next the proposed fault detection scheme is introduced.

3. DECENTRALIZED FAULT DETECTION

Our objective for the fault detection scheme is to design a network of LFDs such that a LFD monitors a subsystem using the local states of that subsystem. Now consider the nonlinear observer for the subsystem 'i' described by

$$\begin{aligned} \dot{\hat{x}}_{ij}(t) &= -\lambda \hat{x}_{ij}(t) + x_{i(j+1)}(t) + \lambda x_{ij}(t) \qquad j = 1, 2, ..., n_i - 1 \\ \dot{\hat{x}}_{in_i}(t) &= -\lambda \hat{x}_{in_i}(t) + f_i\left(x_i(t), u_i(t)\right) + \lambda x_{in_i}(t) + \hat{h}_i\left(x_i(t); \hat{\theta}_i(t)\right) \end{aligned} \tag{2}$$

where $\hat{x}_i(t) \in \mathbb{R}^{n_i}$ is the estimated local state vector of the i[th] subsystem, $\hat{h}_i: \mathbb{R}^{n_i} \times \mathbb{R}^{p_i} \to \mathbb{R}$ is the output of the online approximator with $\hat{\theta}_i \in \mathbb{R}^{p_i}$ being the set of unknown parameters, and λ is a user defined scalar. Initial values of the LFD are taken as $\hat{x}_i(0) = \hat{x}_{i_0}$, $\hat{\theta}_i(0) = \hat{\theta}_{i_0}$, such that $\hat{h}_i(x_i, \hat{\theta}_{i_0}) = 0, \ \forall x_i \in X_i$.

Define a local detection residual $e_i = x_i - \hat{x}_i$. Under healthy system operation, the local residual dynamics are described by

$$\begin{aligned} \dot{e}_{ij}(t) &= -\lambda e_{ij}(t) \qquad j = 1, 2, ..., n_i - 1 \\ \dot{e}_{in_i}(t) &= -\lambda e_{in_i}(t) + g_i\left(x_i(t), \bar{x}_i(t)\right) + \eta_i\left(x(t)\right) \end{aligned}$$

Since the system states are bounded in the healthy operation, the interconnection terms $g_i(x_i(t), \bar{x}_i(t))$ are bounded ($g_i(x(t)) \le g_{i_M}$ in healthy conditions). The uncertainty term $\eta_i(x(t))$ is also bounded by Assumption 1. So by appropriate selection of λ, the local detection residual will remain bounded prior to the occurrence of a fault.

In this work, neural networks (NNs) are used as online approximators, which are off prior to the detection of a fault. When the local detection residual of the k[th] subsystem

exceeds its detection threshold, ρ_k, a fault is declared active in that particular subsystem, and the local OLA that generates $\hat{h}_k(.)$ is initiated. Upon detection, the unknown parameter vector of the k^{th} OLA will be tuned online using the following update law

$$\dot{\hat{\theta}}_k(t) = H\left[\left\|e_k(t)\right\| - \rho_k\right]\left(\alpha_k e_{kn_k}(t)\phi_k(x_k(t)) - \gamma_k\hat{\theta}_k(t)\right) \tag{3}$$

where $\alpha_k > 0$ is the NN learning rate, $0 < \gamma_k < 1$, $\phi_k(x_k(t))$ is the basis function such as sigmoid or RBF, and $H(x)$ is the Heaviside operator defined by

$$H(x) = \begin{cases} 1 & if\ x \geq 0 \\ 0 & if\ x < 0 \end{cases}$$

Then, the output of the OLA is calculated by

$$\hat{h}_k(t) = \hat{\theta}_k^T(t)\phi_k\left(x_k(t)\right)$$

To analytically determine thresholds, consider the solution to the residual dynamics in healthy operating conditions by assuming that the initial values of estimator are equal to those of the actual system

$$e_i(t) = B_i\int_0^t e^{-\lambda(t-\tau)}\left(g_i(x(\tau)) + \eta_i(x_i(\tau))\right)d\tau$$

where $B_i \in \mathbb{R}^{n_i}$ is defined by $B_i = [0,\ldots,0,1]^T$. By taking Frobenius norm and using result of Assumption 1, we will have

$$\left\|e_i(t)\right\| < \int_0^t \left|e^{-\lambda(t-\tau)}\right|\left(\eta_{i_M} + g_{i_M}\right)d\tau$$

where g_{i_M} is the upper bound on interconnection term $g_i(x(t))$ in healthy conditions.

Therefore, no fault will be detected as long as the system is working in healthy operating conditions, if the detection threshold is selected as $\rho_i = (\eta_{i_M} + g_{i_M})(1 - e{-}\lambda t/\lambda$ or $\rho i{=}\eta iM{+}giM/\lambda$.

Upon detection, the local detection residual dynamics of subsystem *k* becomes

$$\dot{e}_{kj}(t) = -\lambda e_{kj}(t) \qquad j = 1,2,\ldots,n_k - 1$$

$$\dot{e}_{kn_k}(t) = -\lambda e_{kn_k}(t) + g_k\left(x_k(t), \bar{x}_k(t)\right) + \eta_k\left(x(t)\right) + h_k\left(x_k(t)\right) - \hat{h}_k\left(x_k(t); \hat{\theta}_k(t)\right)$$

The NLIP assumption on the fault function is asserted to get

$$\dot{e}_{kn_k}(t) = -\lambda e_{kn_k}(t) + g_k\left(x_k(t), \bar{x}_k(t)\right) + \eta_k\left(x(t)\right) + \tilde{\theta}_k^T(t)\phi_k\left(x_k(t)\right) + \varepsilon_k(t) \tag{4}$$

where $\tilde{\theta}_k(t) = \theta_k - \hat{\theta}_k(t)$ is the parameter estimation error and $\varepsilon_k(t)$ is the OLA approximation error. Upon detection of the fault in subsystem '*k*', other subsystems are notified and their fault detection OLADs are not tuned online.

Consequently, for the rest of the subsystems $\tilde{\theta}_i(t) = 0$ and $\varepsilon_i(t) = 0$, and their local residual dynamics are given by

$$\dot{e}_{in_i}(t) = -\lambda e_{in_i}(t) + g_i\left(x_i(t), \bar{x}_i(t)\right) + \eta_i\left(x(t)\right) \qquad for\ i = 1,...,N\ ,\ i \neq k \tag{5}$$

Next fault detectability condition and detection scheme performance are introduced.

Theorem 1 (Fault detectability): Consider the nonlinear subsystem defined by (1) and the local fault detector in (2). A fault in subsystem *k* will be detected, if there exists a time instant, t_d, such that

$$\left| \int_{t_0}^{t_d} e^{-\lambda(t_d - \tau)} \left[g_k\left(x(\tau)\right) + h_k\left(x_k(\tau)\right) \right] d\tau \right| \geq 2\rho_k$$

Proof: The local residual dynamics after the occurrence of fault and prior to the detection is defined by

$$\dot{e}_k(t) = -\lambda e_k(t) + B_k \left[g_k\left(x_k(t), \bar{x}_k(t)\right) + \eta_k\left(x(t)\right) + h_k\left(x_k(t)\right) \right]$$

If the fault occurs at time, t_0, then the solution to the above residual dynamics at time $t_d > t_0$ is given by

$$\begin{aligned} e_k(t_d) &= B_k \int_0^{t_d} e^{-\lambda(t_d - \tau)} \left(g_k\left(x(\tau)\right) + \eta_k\left(x_k(\tau)\right) + h_k\left(x(\tau)\right) \right) d\tau \\ &= B_k \int_0^{t_0} e^{-\lambda(t_d - \tau)} \left(g_k\left(x(\tau)\right) + \eta_k\left(x_k(\tau)\right) \right) d\tau \\ &\quad + B_k \int_{t_0}^{t_d} e^{-\lambda(t_d - \tau)} \left(g_k\left(x(\tau)\right) + \eta_k\left(x_k(\tau)\right) + h_k\left(x(\tau)\right) \right) d\tau \end{aligned}$$

By taking the Frobenius norm, applying Assumption 1, and using the definition of detection residual, it is shown that

$$\begin{aligned} \left\| e_k(t_d) \right\| &\geq \left| \int_{t_0}^{t_d} e^{-\lambda(t_d - \tau)} \left(g_k\left(x(\tau)\right) + h_k\left(x(\tau)\right) \right) d\tau \right| - \int_0^{t_0} e^{-\lambda(t_d - \tau)} \left(g_{k_M} + \eta_{k_M} \right) d\tau - \int_{t_0}^{t_d} e^{-\lambda(t_d - \tau)} \left(\eta_{k_M} \right) d\tau \\ &\geq \left| \int_{t_0}^{t_d} e^{-\lambda(t_d - \tau)} \left(g_k\left(x(\tau)\right) + h_k\left(x(\tau)\right) \right) d\tau \right| - \rho_k \end{aligned}$$

Thus, the fault will be detected if

$$\left| \int_{t_0}^{t_d} e^{-\lambda(t_d-\tau)} \left(g_k\left(x(\tau)\right) + h_k\left(x(\tau)\right) \right) d\tau \right| - \rho_k \geq \rho_k$$

which is equivalent to the given detectability condition.

Note that, in the faulty subsystem, the magnitude of local fault function $h_k\left(x_k(t)\right)$ grows and at some point (t_d) it can satisfy the detectability condition. However, there is no local fault function in other subsystems, but a fault can still be detected in those subsystems if the magnitude of their interconnection term increase due to the nonlocal fault, and satisfy the detectability condition.

Next the following assumption on the interconnection terms is needed before the stability of the proposed observer is discussed upon detection.

Assumption 3: The interconnection terms are unknown and expressed as a function of detection residuals $g_i\left(x(t)\right) \leq \xi_{i0} + \sum_{j=1}^{N} \xi_{ij}\left(e_j(t)\right), i = 1,2,\dots,N,$ where $\xi_{i0}\ for\ i = 1,2,\dots,N$ and $\xi_{ij}(.)\ \ for\ i,j = 1,2,\dots,N$ are unknown constants and smooth functions respectively such that $\xi_{ij}(0) = 0.$

Remark 2: Assumption 3 has been used in a variety of forms in the decentralized control or fault detection literature [14, 15]. Although after the detection of faults, one of the residuals, namely e_k, can go to zero or near zero because of the approximation property of the OLA, other residuals do not, because OLAs in their subsystems are offline. Therefore the bound on the interconnection terms defined in Assumption 3 will hold upon fault detection.

Theorem 2 (Fault Detection Observer Performance): Let the LFD observer network defined in (2) be used to monitor the overall system described by (1), with the local OLA being turned on upon the detection of a fault in the k^{th} subsystem. Let the update law in (3) be used to update the unknown parameter vector$\hat{\theta}_k(t)$. In the presence of system uncertainties, the local FD residuals, $e_i(t)\ \ i = 1,\dots,N,$ and the parameter estimation error, $\tilde{\theta}_k(t)$, will be uniformly ultimately bounded (UUB), provided the design parameters are selected as

$$\lambda > \max_{1 \le i \le N} \frac{3 + N + \xi_{i_{\max}}^2}{2} \tag{6}$$

Proof: Consider the Lyapunov function candidate

$$V(t) = \sum_{i=1}^{N} V_i(t)$$

where $V_i(t) = \frac{1}{2} e_i^T(t) e_i(t) + \frac{1}{2\alpha_i} \tilde{\theta}_i^T(t) \tilde{\theta}_i(t)$. Since $\tilde{\theta}_i(t) = 0 \; for \; i \neq k$, the first derivative of the Lyapunov function is given by

$$\dot{V}(t) = \sum_{i=1}^{N} \left\{ e_i^T(t) \dot{e}_i(t) \right\} + \frac{1}{\alpha_k} \tilde{\theta}_k^T(t) \dot{\tilde{\theta}}_k(t)$$

Substitute $\dot{e}_i(t)$ from residual dynamics in (4) and (5) and $\dot{\tilde{\theta}}_k(t) = -\dot{\hat{\theta}}_k(t)$ from the parameter update law in $\dot{V}$ to get

$$\dot{V}(t) = \sum_{i=1}^{N} \left\{ -\lambda \sum_{j=1}^{n_i - 1} e_{ij}^2(t) - \lambda e_{in_i}^2(t) + e_{in_i}(t) g_i\left(x_i(t), \bar{x}_i(t)\right) + e_{in_i}(t) \eta_i\left(x(t)\right) \right\}$$
$$+ e_{kn_k}(t) \tilde{\theta}_k^T(t) \phi_k\left(x_k(t)\right) + e_{kn_k}(t) \varepsilon_k(t) + \frac{1}{\alpha_k} \tilde{\theta}_k^T(t) \left(-\alpha_k e_{kn_k}(t) \phi_k\left(x_k(t)\right) + \gamma_k \hat{\theta}_k(t) \right)$$

Now apply the Cauchy-Schwarz inequality to arrive at

$$\dot{V}(t) \le \sum_{i=1}^{N} \left\{ -\lambda \sum_{j=1}^{n_i - 1} e_{ij}^2(t) - \lambda e_{in_i}^2(t) + e_{in_i}(t) g_i\left(x_i(t), \bar{x}_i(t)\right) + \frac{e_{in_i}^2(t)}{2} + \frac{\eta_i^2\left(x(t)\right)}{2} \right\}$$
$$+ \frac{e_{kn_k}^2(t)}{2} + \frac{\varepsilon_k^2(t)}{2} + \frac{\gamma_k}{\alpha_k} \tilde{\theta}_k^T(t) \left(\theta_k - \tilde{\theta}_k(t) \right)$$

Assumption 3 and Cauchy-Schwarz inequality are utilized to get

$$\dot{V}(t) \le \sum_{i=1}^{N} \left\{ -\lambda \sum_{j=1}^{n_i - 1} e_{ij}^2(t) - \lambda e_{in_i}^2(t) + \frac{e_{in_i}^2(t)}{2} + \frac{\eta_i^2\left(x(t)\right)}{2} + \frac{\xi_{i0}^2}{2} + \frac{e_{in_i}^2(t)}{2} + \sum_{j=1}^{N} \frac{e_{in_i}^2(t)}{2} + \frac{1}{2} \sum_{j=1}^{N} \xi_{ij}^2\left(e_j(t)\right) \right\}$$
$$+ \frac{e_{kn_k}^2(t)}{2} + \frac{\varepsilon_k^2(t)}{2} - \frac{\gamma_k}{\alpha_k} \left\| \tilde{\theta}_k(t) \right\|^2 + \frac{\gamma_k}{2\alpha_k} \left(\left\| \tilde{\theta}_k(t) \right\|^2 + \left\| \theta_k \right\|^2 \right)$$

After manipulating, the following result is obtained

$$\dot{V}(t) \le \sum_{i=1}^{N}\left\{-\lambda\sum_{j=1}^{n_i-1}e_{ij}^2(t)-\lambda e_{in_i}^2(t)+e_{in_i}^2(t)+\frac{Ne_{in_i}^2(t)}{2}+\frac{\xi_{i0}^2}{2}+\frac{1}{2}\sum_{j=1}^{N}\xi_{ji}^2\left(e_i\left(t\right)\right)+\frac{\eta_{i_M}^2}{2}\right\}$$
$$+\frac{e_{kn_k}^2(t)}{2}+\frac{\varepsilon_{k_M}^2}{2}-\frac{\gamma_k}{2\alpha_k}\left\|\tilde{\theta}_k(t)\right\|^2+\frac{\gamma_k}{2\alpha_k}\left\|\theta_k\right\|^2$$
$$\le \sum_{i=1}^{N}\left\{-\lambda\sum_{j=1}^{n_i-1}e_{ij}^2(t)-\left(\lambda-\frac{2+N}{2}\right)e_{in_i}^2(t)+\frac{1}{2}\sum_{j=1}^{N}\xi_{ji_{\max}}^2\left\|e_i(t)\right\|^2\right\}$$
$$+\sum_{i=1}^{N}\left\{\frac{\eta_{i_M}^2+\xi_{i0}^2}{2}\right\}+\frac{e_{kn_k}^2(t)}{2}+\frac{\varepsilon_{k_M}^2}{2}-\frac{\gamma_k}{2\alpha_k}\left\|\tilde{\theta}_k(t)\right\|^2+\frac{\gamma_k}{2\alpha_k}\left\|\theta_k\right\|^2$$
$$\le \sum_{i=1}^{N}\left\{-\left(\lambda-\frac{2+N}{2}\right)\sum_{j=1}^{n_i}e_{ij}^2(t)+\frac{1}{2}\xi_{i_{\max}}^2\left\|e_i(t)\right\|^2\right\}-\frac{\gamma_k}{2\alpha_k}\left\|\tilde{\theta}_k(t)\right\|^2$$
$$+\sum_{i=1}^{N}\left\{\frac{\eta_{i_M}^2+\xi_{i0}^2}{2}\right\}+\frac{e_{kn_k}^2(t)}{2}+\frac{\varepsilon_{k_M}^2}{2}+\frac{\gamma_k}{2\alpha_k}\left\|\theta_k\right\|^2$$

where $\xi_{i_{\max}}^2=\sum_{j=1}^{N}\xi_{ji_{\max}}^2$. Then, the derivative of the Lyapunov function can be rewritten as

$$\dot{V}(t) \le -\sum_{\substack{i=1\\ i\neq k}}^{N}\left\{\left(\lambda-\frac{2+N+\xi_{i_{\max}}^2}{2}\right)\left\|e_i(t)\right\|^2\right\}-\left(\lambda-\frac{3+N+\xi_{k_{\max}}^2}{2}\right)\left\|e_k(t)\right\|^2-\frac{\gamma_k}{2\alpha_k}\left\|\tilde{\theta}_k(t)\right\|^2$$
$$+\sum_{i=1}^{N}\left\{\frac{\eta_{i_M}^2+\xi_{i0}^2}{2}\right\}+\frac{\varepsilon_{k_M}^2}{2}+\frac{\gamma_k}{2\alpha_k}\left\|\theta_k\right\|^2$$

Therefore, if (6) is satisfied, the derivative of the Lyapunov function will less than zero when

$$\left\|e_i(t)\right\| > \sqrt{\frac{D_M}{\lambda-\frac{2+N+\xi_{i_{\max}}^2}{2}}} \quad i=1,...,N\ ,i\neq k \quad or$$

$$\left\|e_k(t)\right\| > \sqrt{\frac{D_M}{\lambda-\frac{3+N+\xi_{k_{\max}}^2}{2}}} \quad or \quad \left\|\tilde{\theta}_k(t)\right\| > \sqrt{\frac{2\alpha_k D_M}{\gamma_k}}$$

where $D_M=\sum_{i=1}^{N}\frac{\eta_{i_M}^2}{2}+\sum_{i=1}^{N}\frac{\xi_{i0}^2}{2}+\frac{\varepsilon_{k_M}^2}{2}+\frac{\gamma_k}{2\alpha_k}\left\|\theta_k\right\|^2$.

So the local detection residuals and the parameter estimation errors of the online OLA are UUB with the bounds given above.

Remark 3: The bound on e_k is separated from all the other residuals because the bound found for e_k is larger than the bound on other residuals. However the bound on all the residuals, including e_k, can be presented in one condition by taking the supremum among the bounds.

Remark 4: The fault location is identified by communicating the fault detection time at each subsystem to a centralized unit which then compares and finds the minimum time when the fault is detected and its associated subsystem. Then this subsystem will become the fault location. Fault isolation will be dealt in the future.

4. DECENTRALIZED FAULT ACCOMMODATION

Upon determination of the fault location, fault accommodation is performed to mitigate the effect of the fault both in the faulty subsystems and in the other subsystems, because the fault in one subsystem can affect the others through the interconnection terms. The main objective of fault accommodation is to alter the control inputs in order to keep all the system states track their desired trajectories even after the occurrence of an incipient fault. This accommodation is performed in all the subsystems including the one where a fault is detected. As mentioned before, the subsystem where a fault occurs will have two OLAs one for approximating the fault dynamics while the other for accommodation whereas the other subsystems each will have one OLA for accommodation.

First a suitable control input in the healthy conditions is defined, when the ith subsystem dynamics is described by

$$\dot{x}_{ij}(t) = x_{i(j+1)}(t) \qquad j = 1, 2, ..., n_i - 1$$
$$\dot{x}_{in_i}(t) = f_i\left(x_i(t), u_i(t)\right) + g_i\left(x_i(t), \bar{x}_i(t)\right) + \eta_i\left(x(t)\right)$$

Using the technique introduced in [16] this can be rewritten as

$$\dot{x}_{ij}(t) = x_{i(j+1)}(t) \qquad j = 1, 2, ..., n_i - 1$$
$$\dot{x}_{in_i}(t) = f_{i1}\left(x_i(t)\right) + f_{i2}\left(x_i(t)\right)u_i(t) + g_i\left(x_i(t), \bar{x}_i(t)\right) + \eta_i\left(x(t)\right)$$

where $f_{i1}(.)$ and $f_{i2}(.)$ are known smooth functions. If the control objective is to make $x_{i1}(t)$ track the desired trajectory, $x_{id}(t)$, the tracking error dynamics in healthy operating condition, is described by

$$\dot{\bar{e}}_{ij}(t) = \bar{e}_{i(j+1)}(t) \qquad j = 1, 2, ..., n_i - 1$$

$$\dot{\bar{e}}_{in_i}(t) = \dot{x}_{in_i}(t) - x_{id}^{(n_i)}(t)$$

$$= f_{i1}(x_i(t)) + f_{i2}(x_i(t))u_i(t) + g_i(x_i(t), \bar{x}_i(t)) + \eta_i(x(t)) - x_{id}^{(n_i)}(t)$$

where $\bar{e}_i$ represents the vector of local tracking errors in the i[th] subsystem. Now let A_i and B_i be defined as

$$A_i = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix} \qquad B_i = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

Then tracking error dynamics can be rewritten in the matrix form

$$\dot{\bar{e}}_i(t) = A_i\bar{e}_i(t) + B_i\left[f_{i1}(x_i(t)) + f_{i2}(x_i(t))u_i(t) + g_i(x_i(t), \bar{x}_i(t)) + \eta_i(x(t)) - x_{id}^{(n_i)}(t) \right]$$

Since $g_i(x_i(t), \bar{x}_i(t))$ and $\eta_i(x(t))$ are both bounded in healthy operating conditions, the control input $u_i(t) = u_{id}(t)$ defined by

$$u_{id}(t) = f_{i2}(x_i(t))^{-1}\left(x_{id}^{(n_i)}(t) - K_i^T\bar{e}_i(t) - f_{i1}(x_i(t)) \right)$$

can keep tracking error bounded if $K_i \in \mathbb{R}^{n_i}$ is selected such that $A_i - B_iK_i^T$ is Hurwitz.

With the presence of fault, the control input $u_{id}(t)$ can no longer satisfy the control objective. Thus after detection, the local control input is selected as $u_i(t) = u_{id}(t) + u_{ic}(t)$, where $u_{ic}(t)$ is the augmented term to keep the local states track their desired trajectories after the fault. Using this augmented input, the tracking error dynamics can be represented as

$$\dot{\bar{e}}_i(t) = A_i\bar{e}_i(t)$$

$$+B_i\left[-K_i^T\bar{e}_i(t) + f_{i2}(x_i(t))u_{ic}(t) + \eta_i(x(t)) + g_i(x_i(t), \bar{x}_i(t)) + \Omega_i(t - t_0)h_i(x_i(t)) \right]$$

Ideally, $u_{ic}(t)$ should be selected as

$$u_{ic}(t) = f_{i2}(x_i(t))^{-1}\left(g_i(x_i(t), \bar{x}_i(t)) + \Omega_i(t - t_0)h_i(x_i(t)) \right)$$

However, since the fault function and interconnection term are unknown, $u_{ic}(t)$ cannot be practically determined this way. Therefore the output of detection OLA $\hat{h}_i\left(x_i(t);\hat{\theta}_i(t)\right)$ which is the estimation of fault, along with another online approximator referred to as fault accommodation OLA which is used to compensate for the interconnection effects, are utilized to construct the estimated $u_{ic}(t)$ as follows

$$\hat{u}_{ic}(t)=f_{i2}\left(x_i(t)\right)^{-1}\left(-\hat{q}_i\left(\bar{e}_i(t);\hat{W}_i(t)\right)-\hat{h}_i\left(x_i(t);\hat{\theta}_i(t)\right)\right)+f_{i2}\left(x_i(t)\right)^{-1}\left(-l_iB_i^TP_i\bar{e}_i(t)\right) \quad (7)$$

where l_i is a positive constant defined by the user and $P_i>0$ is obtained from the Lyapunov equation $P_i(A_i-B_iK_i^T)+(A_i-B_iK_i^T)^TP_i=-Q_i$ for any positive definite matrix Q_i. Moreover $\hat{q}_i\left(\bar{e}_i(t);\widehat{W}_i(t)\right)$ is defined as the output of the online approximator from accommodation $\hat{q}_i\left(\bar{e}_i(t);\hat{W}_i(t)\right)=\hat{W}_i^T(t)\psi_i\left(\bar{e}_i(t)\right)$ which is turned on upon detection of a fault, $\psi_i\left(\bar{e}_i(t)\right)$ is the basis function and $\widehat{W}_i$ is the estimated parameter vector which will be updated by an adaptive update law (to be defined later) in order to ensure the stability of the accommodation scheme.

Note that $h_i\left(x_i(t)\right)=\hat{h}_i\left(x_i(t);\hat{\theta}_i(t)\right)=0$ *if* $i\neq k$. Thus $\tilde{\theta}_i^T(t)\phi_i\left(x_i(t)\right)+\varepsilon_i(t)=\tilde{h}_i=0$ *if* $i\neq k$. Then the tracking error dynamics can be rewritten as

$$\begin{aligned}\dot{\bar{e}}_i(t)=A_i\bar{e}_i(t)+B_i\Big[&-K_i^T\bar{e}_i(t)-B_i^TP_i\bar{e}_i(t)-\hat{W}_i^T(t)\psi_i\left(x_i(t)\right)\\&+\eta_i\left(x(t)\right)+g_i\left(x_i(t),\bar{x}_i(t)\right)+\tilde{\theta}_i^T(t)\phi_i\left(x_i(t)\right)+\varepsilon_i(t)\Big]\end{aligned} \quad (8)$$

Furthermore, combined tracking error is defined as $s_i(t)=B_i^TP_i\bar{e}_i(t)$ for convenience in the proof of next theorem. Next Assumption 4 is introduced which is similar to Assumption 3, before considering the performance of fault accommodation scheme in Theorem 3.

Assumption 4 [14]: The interconnection terms are unknown but expressed as $g_i\left(x(t)\right)\leq\varsigma_{i0}+\sum_{j=1}^{N}\varsigma_{ij}\left(|s_j(t)|\right),\ i=1,2,\dots,N,$ where $\varsigma_{i0}\ for\ i=1,2,\dots,N$ are unknown constants and $\varsigma_{ij}(.)\ for\ i,j=1,2,\dots,N$ are unknown smooth functions, such that $\varsigma_{ij}(0)=0$

Theorem 3 (Performance of the Fault Accommodation Scheme): Consider the large-scale interconnected system described by (1). Upon detecting a fault in the k^{th} subsystem, let the control input of all the subsystems be augmented as $u_i(t) = u_{id}(t) + \hat{u}_{ic}(t)$ where $\hat{u}_{ic}(t)$ is defined in (7) and the parameter update law of the accommodation online approximator is given by

$$\dot{\hat{W}}_i(t) = \beta_i \psi_i(\bar{e}_i(t)) s_i(t) - \upsilon_i \hat{W}_i(t) \tag{9}$$

with υ_i and β_i being positive constants. Then the tracking errors $\bar{e}_i(t)$, and parameter estimation errors $\tilde{W}_i(t) = W_i - \hat{W}_i(t)$ are UUB, if the design parameters are selected such that

$$l_i > 3 \;\; and \;\; \gamma_i > \alpha_i \phi_{i_{\max}}^2 \qquad for \; i = 1, ..., N \tag{10}$$

Proof: Consider the Lyapunov function candidate as $V(t) = \sum_{i=1}^{N} V_i(t)$ where

$$V_i(t) = \frac{1}{2}\bar{e}_i^T(t) P_i \bar{e}_i(t) + \frac{1}{2\beta_i}\tilde{W}_i^T(t)\tilde{W}_i(t) + \frac{1}{2}e_i^T(t)e_i(t) + \frac{1}{2\alpha_i}\tilde{\theta}_i^T(t)\tilde{\theta}_i(t)$$

Then the derivative of the Lyapunov function is given by

$$\dot{V}(t) = \sum_{i=1}^{N}\left\{\frac{1}{2}\left(\dot{\bar{e}}_i^T(t) P_i \bar{e}_i(t) + \bar{e}_i^T(t) P_i \dot{\bar{e}}_i(t)\right) + \frac{1}{\beta_i}\tilde{W}_i^T(t)\dot{\tilde{W}}_i(t)\right\}$$
$$+ \sum_{i=1}^{N}\left\{e_i^T(t)\dot{e}_i(t)\right\} + \frac{1}{\alpha_k}\tilde{\theta}_k^T(t)\dot{\tilde{\theta}}_k(t)$$

And after substituting $\bar{e}_i(k)$ from the tracking error dynamics, it can be described by

$$\dot{V}(t) = \sum_{i=1}^{N}\left\{\frac{1}{2}\bar{e}_i^T(t)\left(A_i^T P_i + P_i A_i\right)\bar{e}_i(t) - \frac{1}{\beta_i}\tilde{W}_i^T(t)\dot{\hat{W}}_i(t)\right.$$
$$\left. + \bar{e}_i^T(t) P_i B_i\left[-K_i \bar{e}_i(t) - l_i B_i^T P_i \bar{e}_i(t) + \eta_i(x(t)) - \hat{W}_i^T(t)\psi_i\left(\bar{e}_i(t)\right) + g_i\left(x_i(t), \bar{x}_i(t)\right)\right]\right\}$$
$$+ \bar{e}_k^T(t) P_k B_k\left(\tilde{\theta}_k^T(t)\phi_k\left(x_k(t)\right) + \varepsilon_k(t)\right) + \sum_{i=1}^{N}\left\{e_i^T(t)\dot{e}_i(t)\right\} + \frac{1}{\alpha_k}\tilde{\theta}_k^T(t)\dot{\tilde{\theta}}_k(t)$$

$$= \sum_{i=1}^{N}\left\{-\frac{1}{2}\bar{e}_i^T(t) Q_i \bar{e}_i(t) - \frac{1}{\beta_i}\tilde{W}_i^T(t)\dot{\hat{W}}_i(t)\right.$$
$$\left. + s_i(t)\left[-l_i s_i(t) - \hat{W}_i^T(t)\psi_i\left(\bar{e}_i(t)\right) + g_i\left(x_i(t), \bar{x}_i(t)\right) + \eta_i(x(t))\right]\right\}$$
$$+ s_k(t)\left(\tilde{\theta}_k^T(t)\phi_k\left(x_k(t)\right) + \varepsilon_k(t)\right) + \sum_{i=1}^{N}\left\{e_i^T(t)\dot{e}_i(t)\right\} + \frac{1}{\alpha_k}\tilde{\theta}_k^T(t)\dot{\tilde{\theta}}_k(t)$$

Now we use the result of assumption 4 and the Cauchy-Schwarz inequality to get

$$\dot{V}(t) \le \sum_{i=1}^{N} \left\{ -\frac{1}{2}\bar{e}_i^T(t) Q_i \bar{e}_i(t) - \frac{1}{\beta_i}\tilde{W}_i^T(t)\dot{\hat{W}}_i(t) - l_i s_i^2(t) - s_i(t)\hat{W}_i^T(t)\psi_i\left(\bar{e}_i(t)\right) \right.$$
$$\left. + \frac{s_i^2(t)}{2} + \frac{N}{2}\sum_{j=1}^{N}\varsigma_{ij}^2\left(\left|s_j(t)\right|\right) + \frac{s_i^2(t)}{2} + \frac{\left(\eta_i(x(t))\right)^2}{2} + \frac{s_i^2(t)}{2} + \frac{\varsigma_{i0}^2}{2} \right\}$$
$$+ \frac{s_k^2(t)}{2} + \frac{\left(\tilde{\theta}_k^T(t)\phi_k\left(x_k(t)\right)\right)^2}{2} + \frac{s_k^2(t)}{2} + \frac{\left(\varepsilon_k(t)\right)^2}{2} + \sum_{i=1}^{N}\left\{e_i^T(t)\dot{e}_i(t)\right\} + \frac{1}{\alpha_k}\tilde{\theta}_k^T(t)\dot{\tilde{\theta}}_k(t)$$

Since $\varsigma_{ij}(\bullet)$ is a smooth function for *i,j=1,…,N*, there exists another smooth function $\zeta_{ij}(\bullet)$ such that $\varsigma_{ij}(|s_j|) = |s_j|\zeta_{ij}(|s_j|)$ for *i,j=1,…,N* [14]. Applying this result to the derivative of the Lyapunov function and changing the order of the summations leads to

$$\dot{V}(t) \le \sum_{i=1}^{N} \left\{ -\frac{1}{2}\bar{e}_i^T(t) Q_i \bar{e}_i(t) - \frac{1}{\beta_i}\tilde{W}_i^T(t)\dot{\hat{W}}_i(t) - \left(l_i - \frac{3}{2}\right)s_i^2(t) - s_i(t)\hat{W}_i^T(t)\psi_i\left(\bar{e}_i(t)\right) \right.$$
$$\left. + \frac{N}{2}\left|s_i(t)\right|^2\sum_{j=1}^{N}\zeta_{ji}^2\left(\left|s_i(t)\right|\right) + \frac{\left(\eta_i(x(t))\right)^2}{2} + \frac{\varsigma_{i0}^2}{2} \right\}$$
$$+ s_k^2(t) + \frac{\left(\tilde{\theta}_k^T(t)\phi_k\left(x_k(t)\right)\right)^2}{2} + \frac{\left(\varepsilon_k(t)\right)^2}{2} + \sum_{i=1}^{N}\left\{e_i^T(t)\dot{e}_i(t)\right\} + \frac{1}{\alpha_k}\tilde{\theta}_k^T(t)\dot{\tilde{\theta}}_k(t)$$

The function $q_i\left(\bar{e}_i(t)\right) = \frac{N}{2}s_i(t)\sum_{j=1}^{N}\zeta_{ji}^2\left(\left|s_i(t)\right|\right)$ is a smooth function, so it can be approximated by using two-layer NN with bounded activation functions, target weight parameters, and estimation error as $q_i\left(\bar{e}_i(t)\right) = W_i^T(t)\psi_i\left(\bar{e}_i(t)\right) + \mu_i(t)$, where $|\mu_i| \le \mu_{i_M}$. Now use $\tilde{W}_i = W_i - \hat{W}_i$ and rewrite the derivative of the Lyapunov function as

$$\dot{V}(t) \le \sum_{i=1}^{N} \left\{ -\frac{1}{2}\bar{e}_i^T(t) Q_i \bar{e}_i(t) - \frac{1}{\beta_i}\tilde{W}_i^T(t)\dot{\hat{W}}_i(t) - \left(l_i - \frac{3}{2}\right)s_i^2(t) \right.$$
$$\left. + s_i(t)\tilde{W}_i^T(t)\psi_i\left(\bar{e}_i(t)\right) + s_i(t)\mu_i(t) + \frac{\varsigma_{i0}^2}{2} + \frac{\left(\eta_i(x(t))\right)^2}{2} \right\}$$
$$+ s_k^2(t) + \frac{\left(\tilde{\theta}_k^T(t)\phi_k\left(x_k(t)\right)\right)^2}{2} + \frac{\left(\varepsilon_k(t)\right)^2}{2} + \sum_{i=1}^{N}\left\{e_i^T(t)\dot{e}_i(t)\right\} + \frac{1}{\alpha_k}\tilde{\theta}_k^T(t)\dot{\tilde{\theta}}_k(t)$$

Taking the same steps as in the proof of theorem 2, the last line in the above inequality can be rewritten as

$$\sum_{i=1}^{N}\left\{e_i^T(t)\dot{e}_i(t)\right\}+\frac{1}{\alpha_k}\tilde{\theta}_k^T(t)\dot{\tilde{\theta}}_k(t)\le-\sum_{\substack{i=1\\ i\neq k}}^{N}\left\{\left(\lambda-\frac{2+N+\xi_{i_{\max}}^2}{2}\right)\left\|e_i(t)\right\|^2\right\}-\left(\lambda-\frac{3+N+\xi_{k_{\max}}^2}{2}\right)\left\|e_k(t)\right\|^2$$
$$-\frac{\gamma_k}{2\alpha_k}\left\|\tilde{\theta}_k(t)\right\|^2+\sum_{i=1}^{N}\left\{\frac{\eta_{i_M}^2+\xi_{i0}^2}{2}\right\}+\frac{\varepsilon_{k_M}^2}{2}+\frac{\gamma_k}{2\alpha_k}\left\|\theta_k\right\|^2$$

If the update law is selected as in (9), then

$$\dot{V}(t)\le\sum_{i=1}^{N}\left\{-\frac{1}{2}\bar{e}_i^T(t)Q_i\bar{e}_i(t)+\frac{\upsilon_i}{\beta_i}\tilde{W}_i^T(t)\hat{W}_i(t)-(l_i-2)s_i^2(t)+\frac{\mu_i^2(t)}{2}+\frac{\varsigma_{i0}^2}{2}+\frac{\left(\eta_i(x(t))\right)^2}{2}\right\}$$
$$+s_k^2(t)+\frac{\left(\tilde{\theta}_k^T(t)\phi_k(x_k(t))\right)^2}{2}+\frac{\left(\varepsilon_k(t)\right)^2}{2}-\sum_{i=1}^{N}\left\{\left(\lambda-\frac{3+N+\xi_{i_{\max}}^2}{2}\right)\left\|e_i(t)\right\|^2\right\}-\frac{\gamma_k}{2\alpha_k}\left\|\tilde{\theta}_k(t)\right\|^2$$
$$+\sum_{i=1}^{N}\left\{\frac{\eta_{i_M}^2+\xi_{i0}^2}{2}\right\}+\frac{\varepsilon_{k_M}^2}{2}+\frac{\gamma_k}{2\alpha_k}\left\|\theta_k\right\|^2$$

$$\le\sum_{i=1}^{N}\left\{-\frac{1}{2}\bar{e}_i^T(t)Q_i\bar{e}_i(t)-\frac{\upsilon_i}{2\beta_i}\left\|\tilde{W}_i(t)\right\|^2-(l_i-3)s_i^2(t)-\left(\lambda-\frac{3+N+\xi_{i_{\max}}^2}{2}\right)\left\|e_i(t)\right\|^2\right\}$$
$$-\left(\frac{\gamma_k-\alpha_k\phi_{k_{\max}}^2}{2\alpha_k}\right)\left\|\tilde{\theta}_k(t)\right\|^2+\sum_{i=1}^{N}\left\{\frac{\upsilon_i}{2\beta_i}\left\|W_i\right\|^2+\frac{\mu_{i_M}^2}{2}+\frac{\varsigma_{i0}^2}{2}+\frac{\xi_{i0}^2}{2}+\eta_{i_M}^2\right\}+\varepsilon_{k_M}^2+\frac{\gamma_k}{2\alpha_k}\left\|\theta_k\right\|^2$$

Assuming that user defined parameters are selected such that $l_i>3$ and $\gamma_i>\alpha_i\phi_{i_{max}}^2$, $\dot{V}(t)$ will be less than zero, if one of the following $3N+1$ conditions are satisfied

$$\left\|\bar{e}_i\right\|>\sqrt{\frac{2\bar{D}_M}{Q_{i_{\max}}}}\qquad i=1,\ldots,N$$

$$\left\|\tilde{W}_i(t)\right\|>\sqrt{\frac{2\beta_i\bar{D}_M}{\upsilon_i}}\qquad i=1,\ldots,N$$

$$\left\|e_i(t)\right\|>\sqrt{\frac{2\bar{D}_M}{2\lambda-3-N-\xi_{i_{\max}}^2}}\qquad i=1,\ldots,N$$

$$\left\|\tilde{\theta}_k(t)\right\|>\sqrt{\frac{2\alpha_k\bar{D}_M}{\gamma_k-\alpha_k\phi_{k_{\max}}^2}}$$

where $\bar{D}_M=\sum_{i=1}^{N}\left\{\frac{\upsilon_i}{2\beta_i}\left\|W_i\right\|^2+\frac{\mu_{i_M}^2}{2}+\frac{\varsigma_{i0}^2}{2}+\frac{\xi_{i0}^2}{2}+\eta_{i_M}^2\right\}+\varepsilon_{k_M}^2+\frac{\gamma_k}{2\alpha_k}\left\|\theta_k\right\|^2$

So all the tracking errors $\bar{e}_i$ and the parameter estimation errors $\widetilde{W}_i$ will be UUB with the bounds provided above. Furthermore, this proof shows the stability of the overall system, since it also guarantees the boundedness of detection residuals e_i and the parameter estimation error $\tilde{\theta}_k$.

Remark 5: Note that the detection OLA approximates the fault function $\hat{h}_k$ (with a bounded error), while the accommodation OLA which is generating $\hat{q}_i$ is not estimating the interconnection functiong_i but it's used to approximate $q_i(\bar{e}_i(t)) = \frac{N}{2}\bar{e}_i^T(t)P_iB_i\sum_{j=1}^{N}\zeta_{ji}^2\left(\left|\bar{e}_i^T(t)P_iB_i\right|\right)$. In fact, all the accommodation OLAs must be utilized together to cancel all the interconnection terms.

Another approach for performing fault accommodation is to apply a single OLA in each subsystem right after detection, and refrain from utilizing the detection OLA. When this approach is chosen, the complexity of the overall scheme is reduced and less processing is required. In this case, the accommodation OLA will estimate$q_i(\bar{e}_i(t)) + h_i(x_i(t))$. Although the output of this OLA is used to cancel the fault effect, it does not have any other benefits since there is no physical interpretation of the function which is being estimated. In addition, the time to failure determination cannot be conducted upon detection.

In contrast, the proposed detection and accommodation scheme which uses two online approximators is more involved since one OLA continues to provide fault approximation while the other is used for accommodation in the faulty subsystem while a fault accommodation OLA in the others. The estimation of fault dynamics can be utilized to perform time-to-failure determination (which has been done in our previous work [17] for nonlinear interconnected discrete-time systems), at the same time when accommodation is being done. It's worth mentioning that although the effect of a fault on the overall performance of the system can be mitigated by the accommodation scheme, the root cause of the fault will not be eliminated and at some point it could result in internal damage to system components. Determination of fault type and time-to-failure will definitely be helpful in preventing such problems.

5. PREDICTION

The stability of tracking error dynamics after initiation of fault accommodation was demonstrated in Theorem 3. However, in practice it is important to know the time that is required by the accommodation scheme to regain the desired tracking performance. More importantly, it is necessary to know if the accommodation goal can be achieved before the system reaches a failure point. This problem can be addressed by using both time-to-failure and time-to-accommodation estimation.

Time-to-failure (TTF) which is also referred to as remaining useful life of the system can be estimated upon detecting a fault, by comparing the estimated parameters of the OLAD to the user defined failure limit. The following theorem, which is introduced in our previous work [18], provides an analytical formula for TTF estimation.

Theorem 4 (TTF Determination) : If the fault location is found to be the i^{th} subsystem, TTF for the j^{th} parameter of the fault, at the time t, can be determined by

$$TTF_j(t) = \frac{1}{\gamma_i} \ln\left(\left| \frac{\gamma_i \hat{\theta}_{i_j}(t) - \alpha_i e_{in_i}(t) \phi_{i_j}(x_i(t))}{\gamma_i \theta_{i_{\text{j max}}} - \alpha_i e_{in_i}(t) \phi_{i_j}(x_i(t))} \right| \right) \tag{11}$$

where $\theta_{i_{\text{j max}}}$ is the failure limit regarding the j^{th} parameter of fault in the i^{th} subsystem, in terms of maximum value of the system parameter , $\theta_{i_{\text{j}}}$.

Proof: Upon detection of fault in the i^{th} subsystem, the parameter update law of the detection OLA in that subsystem is defined by

$$\dot{\hat{\theta}}_i(t) = -\gamma_i \hat{\theta}_i(t) + \alpha_i e_{in_i}(t) \phi_i(x_i(t))$$

Since $\gamma_i > 0$, this state space equation represents a stable system and its solution is given by

$$\hat{\theta}_i(t) = e^{-\gamma_i (t-t_0)} \hat{\theta}_i(t_0) + \int_{t_0}^{t} e^{-\gamma_i (t-\tau)} \alpha_i e_{in_i}(\tau) \phi_i(x_i(\tau)) d\tau$$

We know that the basis function $\phi_i(\cdot)$ is upper bounded and it was shown in Theorem 2, that the detection residual is bounded. Thus, the term $\alpha_i e_{in_i}(\tau) \phi_i(x_i(\tau))$ in

the above equation is bounded and if assumed to be held constant at time *t*. Suppose that t_f is the time at which the system attains failure. Then $\hat{\theta}_i(t_f)$ can be described by

$$\hat{\theta}_i\left(t_f\right) = e^{-\gamma_i\left(t_f - t\right)}\hat{\theta}_i\left(t\right) + \alpha_i e_{in_i}\left(t\right)\phi_i\left(x_i(t)\right)\int_t^{t_f} e^{-\gamma_i\left(t_f-\tau\right)}d\tau$$

$$= e^{-\gamma_i\left(t_f - t\right)}\hat{\theta}_i\left(t\right) + \alpha_i \frac{1-e^{-\gamma_i\left(t_f - t\right)}}{\gamma_i} e_{in_i}\left(t\right)\phi_i\left(x_i(t)\right)$$

Therefore, $\hat{\theta}_{i_j}(t_f)$, which is the j^{th} element of $\hat{\theta}_i(t_f)$, can be represented by

$$\hat{\theta}_{i_j}\left(t_f\right) = \frac{\alpha_i}{\gamma_i} e_{in_i}(t)\phi_{i_j}\left(x_i(t)\right) + e^{-\gamma_i\left(t_f - t\right)}\left(\hat{\theta}_{i_j}\left(t\right) - \frac{\alpha_i}{\gamma_i} e_{in_i}(t)\phi_{i_j}\left(x_i(t)\right)\right)$$

where ϕ_{i_j} is the j^{th} element of the basis function of fault in the i^{th} subsystem. After substituting $\hat{\theta}_{i_j}\left(t_f\right)$ with $\theta_{i_{\text{j max}}}$ we get

$$\frac{\gamma_i\theta_{i_{\text{j max}}} - \alpha_i e_{in_i}(t)\phi_{i_j}\left(x_i(t)\right)}{\gamma_i\hat{\theta}_{i_j}\left(t\right) - \alpha_i e_{in_i}(t)\phi_{i_j}\left(x_i(t)\right)} = e^{-\gamma_i\left(t_f - t\right)}$$

Since t_f is the time of failure and *t* is the current time, $\left(t_f - t\right)$ is the time to failure or TTF. With simple mathematical manipulations we finaly arrive at

$$TTF_j(t) = \frac{1}{\gamma_i}\ln\left(\left|\frac{\gamma_i\hat{\theta}_{i_j}\left(t\right) - \alpha_i e_{in_i}(t)\phi_{i_j}\left(x_i(t)\right)}{\gamma_i\theta_{i_{\text{j max}}} - \alpha_i e_{in_i}(t)\phi_{i_j}\left(x_i(t)\right)}\right|\right)$$

Figure 5.1 shows magnitude of the tracking error in the faulty subsystem. The fault will definitely cause the tracking error to increase, since the nominal controller does not have any information about it. Upon detecting the fault, the accommodation scheme will modify the controller to mitigate fault effects and decrease the tracking error. The problem of time-to-accommodation estimation is then defined by continuously estimating the time left until the tracking error magnitude decreases below a desired limit defined by the user. In the next theorem, linear approximation is used to estimate time-to-accommodation (TTA) online.

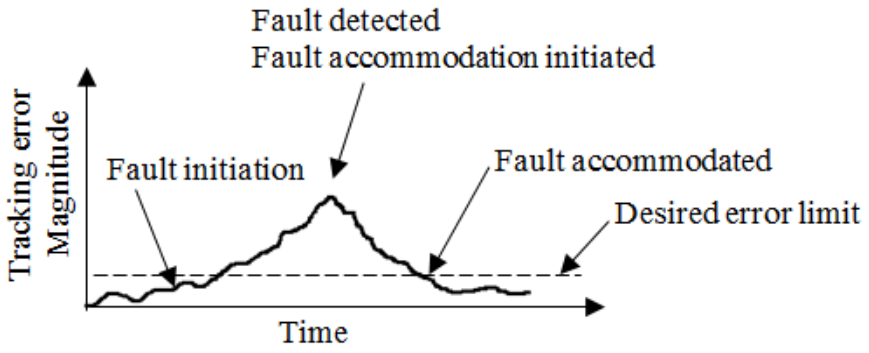

Figure 5.1. Tracking error with fault accommodation

Theorem 5: Upon detecting a fault in the subsystem "*i*" and initiation of fault accommodation scheme, time-to-accommodation can be estimated online by using

$$TTA_i(t) = \min_{1 \le i \le n_i} \left| \frac{\sigma_{ij} - \left|\bar{e}_{ij}(t)\right|}{\dot{\bar{e}}_{ij}(t)} \right| \tag{12}$$

where $\sigma_{ij} > 0$ is the desired tracking error limit for the jth state in subsystem *i*.

Proof: By using linear approximation time-to-accommodation for each of the states at time *t* can be found by

$$TTA_{ij}(t) = \left| \frac{\sigma_{ij} - \left|\bar{e}_{ij}(t)\right|}{\dot{\bar{e}}_{ij}(t)} \right| \quad for \;\; j = 1, \ldots, n_i$$

Since the tracking error for all of the states must be less than their limits, the total TTA is obtained as the maximum among all the individual time-to-accommodations, i.e.

$$TTA_i(t) = \min_{1 \le i \le n_i} \left| \frac{\sigma_{ij} - \left|\bar{e}_{ij}(t)\right|}{\dot{\bar{e}}_{ij}(t)} \right|$$

Theorems 4 and 5 provide analytical formulas for online estimation of TTF and TTA. This can be utilized to determine whether or not the accommodation scheme can bring the system performance back to normal before the complete failure of the system. Figure 5.2 illustrates the process of detection and accommodation. Upon detecting a fault, the accommodation unit is activated and TTF and TTA are activated. Then estimated

TTF and TTA are compared to each other. If either the TTF decreases below a certain limit (T_L) or TTF is smaller than TTA, then the system is shut down for maintenance, since the accommodation unit cannot keep the system in safe operating conditions. If both of these situations do not happen, the accommodation scheme and TTF/TTA estimation continues and the system will keep running uninterruptedly.

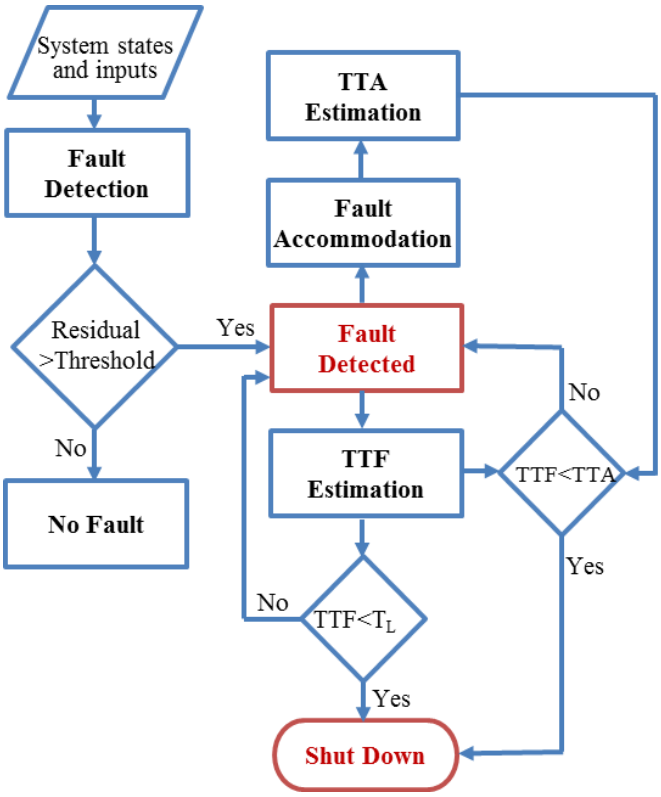

Figure 5.2. Detection and accommodation flow chart

6. SIMULATION RESULTS

A system of double inverted pendulums [14], which is depicted in Figure 6.1 [19], is used to verify the proposed decentralized detection and accommodation schemes. The pendulums are connected to each other by a spring and their motion dynamics are described by

$$\dot{\theta}_1 = \omega_1$$

$$\dot{\omega}_1 = \frac{m_1 g l_1 sin\theta_1 - b_1\omega_1}{m_1 l_1^2} + \frac{u_1}{m_1 l_1^2} + \frac{Fa_1 \cos(\theta_1 - \beta)}{m_1 l_1^2}$$

$$\dot{\theta}_2 = \omega_2$$

$$\dot{\omega}_2 = \frac{m_2 g l_2 sin\theta_2 - b_2\omega_2}{m_2 l_2^2} + \frac{u_2}{m_2 l_2^2} - \frac{Fa_2 \cos(\theta_2 - \beta)}{m_2 l_2^2}$$

where $\theta_1 = x_{11}$ and $\theta_2 = x_{21}$ are the angular displacements of the pendulums from the vertical position, $\omega_1 = x_{12}$ and $\omega_2 = x_{22}$ are the angular velocities of the pendulums, b_1 and b_2 are damping coefficients, g is standard gravity, m_1 and m_2 are masses and l_1 and l_2 are lengths of the pendulums. Moreover

$$F = k\left(1 + A^2\left(l_k - l_0\right)^2\right)(l_k - l_0)$$

$$\beta = arctan\left(\frac{a_1 cos\theta_1 - a_2 cos\theta_2}{l_0 - a_1 sin\theta_1 + a_2 sin\theta_2}\right)$$

$$l_k = \sqrt{\left(l_0 + a_2 sin\theta_1 - a_1 sin\theta_1\right)^2 + \left(a_1 cos\theta_1 - a_2 cos\theta_2\right)^2}$$

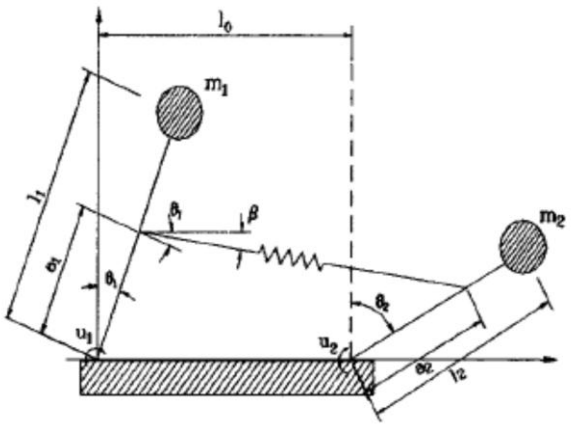

Figure 6.1. Double inverted pendulums

Uncertainties in the form of $\eta_i(x(t)) = 10^{-2}\cos(x_{i1}(t))$ are added to the dynamics of the system and a fault is seeded in the first pendulum subsystem at time $t_0 = 10\ sec$. The fault is described by

$$\Omega_1(t-t_0)h_1(x_1(t)) = \begin{cases} 0, & if\ t < t_0 \\ (1-e^{-0.1(t-t_0)})\dfrac{x_{11}(t)+10}{m_1 l_1^2}, & if\ t \geq t_0 \end{cases}$$

The values of system parameters are provided as $a_1 = 0.1$, $a_2 = 0.35$, $b_1 = b_2 = 0.009$, $m_1 = 0.4$, $m_2 = 0.5$, $l_1 = 0.3$, $l_2 = 0.35$, $l_0 = 0.4$, $k = 30$, and $A = 0.1$. The initial states are $[\theta_1(0), \omega_1(0)]^T = [\theta_2(0), \omega_2(0)]^T = [0.2, 0]^T$, the desired trajectories are $x_{1d}(t) = \theta_{1d}(t) = 0.55\cos(3t)$ and $x_{2d}(t) = \theta_{2d}(t) = 0.35\cos(4t)$ and the simulation runs for 50 seconds. The control inputs in healthy conditions are

$$u_{1d}(t) = -m_1 g l_1 \sin x_{11}(t) + b_1 x_{12}(t) + m_1 l_1^2 \left(\ddot{x}_{1d}(t) - K^T \bar{e}_1(t)\right)$$
$$u_{2d}(t) = -m_2 g l_2 \sin x_{21}(t) + b_2 x_{22}(t) + m_2 l_2^2 \left(\ddot{x}_{2d}(t) - K^T \bar{e}_2(t)\right)$$

where $\bar{e}_i$ is the tracking error vector defined by $\bar{e}_i = [x_{i1} - x_{id},\ x_{i2} - \dot{x}_{id}]^T$ and $K = [100, 1]^T$

To perform the decentralized fault detection, the local fault detectors in (2) are utilized. Upon detection the parameter update law in (3) is used to update the unknown parameter vector such that the online approximator in the first LFD estimates the fault function. The LFD parameters are selected as $\lambda = 10$, $\rho = 5$, $\alpha_1 = 50$, and $\gamma_1 = 0.001$. Also after the detection, decentralized fault accommodation is performed and control inputs are augmented as $u_i = u_{id} + \hat{u}_{ic}$ to cancel the fault effects. $\hat{u}_{ic}(t)$ is calculated by

$$\hat{u}_{ic}(t) = m_i l_i^2 \left(-\hat{q}_i\left(\bar{e}_i(t); \hat{W}_i(t)\right) - \hat{h}_i\left(x_i(t); \hat{\theta}_i(t)\right)\right) + m_i l_i^2 \left(-c_i B_i^T P_i \bar{e}_i(t)\right) \qquad i = 1, 2$$

where $\hat{h}_i$ is the output of online approximator in ith LFD, $\hat{q}_i$ is the output of accommodation OLA in ith subsystem, $B_1 = B_2 = [0\ \ 1]^T$, $c_1 = c_2 = 50$, and P_1 and P_2 are obtained from solving the Lyapunov equation $P_i(A_i - B_i K_i^T) + (A_i - B_i K_i^T)^T P_i = -Q_i$. By choosing $Q_1 = Q_2 = I_{2x2}$ we get $P_1 = P_2 = \begin{bmatrix} 50.505 & 0.005 \\ 0.005 & 50.505 \end{bmatrix}$.

The unknown parameter vector $\widehat{W}_i$is tuned online using update law in (8) with $\beta_1 = \beta_2 = 50$ and $\upsilon_1 = \upsilon_2 = 0.001$. It should also be mentioned that neural networks with five hidden layer neurons are used as online approximators.

Figure 6.2 shows the detection residual of the first pendulum (subsystem 1) along with the detection threshold. The residual is below the threshold in the healthy operating conditions. When the fault occurs at $t_0 = 10\ sec$, this residual starts to increase and

eventually it exceeds the detection threshold at $t = 12.65\ sec$. At this point the fault is detected in the first subsystem, the online approximator is turned on to approximate the fault dynamics, and the parameter update law in (3) is used to update the unknown parameters. As observed in Figure 6.3, the online approximator is able to estimate the fault function with a small error, after the fault is detected. Because of this approximation property of the OLA, the estimator is able to adapt with the changes in the actual system due to the fault and the detection residual falls below the threshold again.

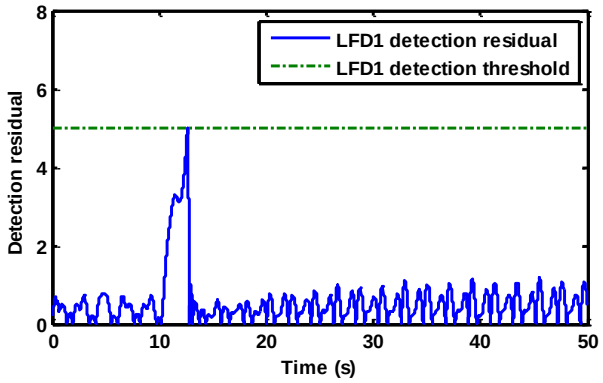

Figure 6.2. Subsystem 1 detection residual and threshold

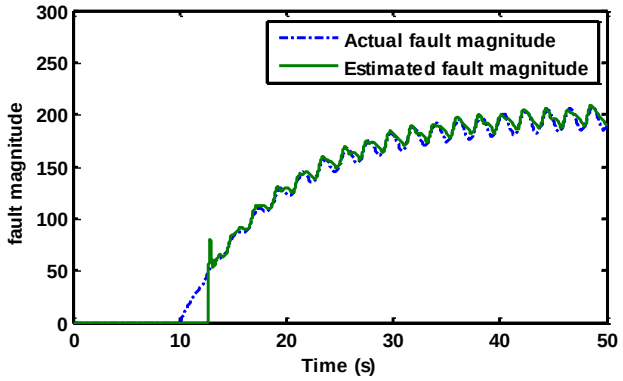

Figure 6.3. Actual and estimated fault functions

Figure 6.4 depicts the tracking error if no fault accommodation is performed. In this case the fault causes the tracking error to increase. In contrast, when the proposed decentralized fault accommodation strategy is adopted, the estimated fault function and

another neural network are used to cancel the fault effects. It can be observed in Figure 6.5 that the tracking error of pendulum 1 is brought back close to zero upon detection of the fault, when the accommodation scheme is applied. The effectiveness of the proposed scheme, which does not require transmission of actual or estimated states among subsystems, can be concluded by comparing the tracking errors in Figures 6.4 and 6.5. The estimation of time-to-accommodation is shown in Figure 6.6. The accommodation scheme requires about 0.06 seconds to bring the system in desired operating condition.

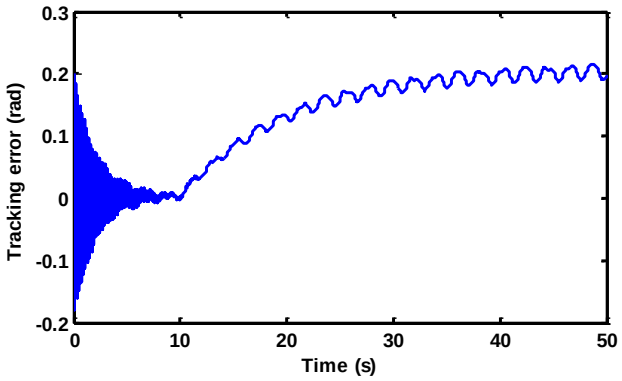

Figure 6.4. Pendulum 1 tracking error without accommodation

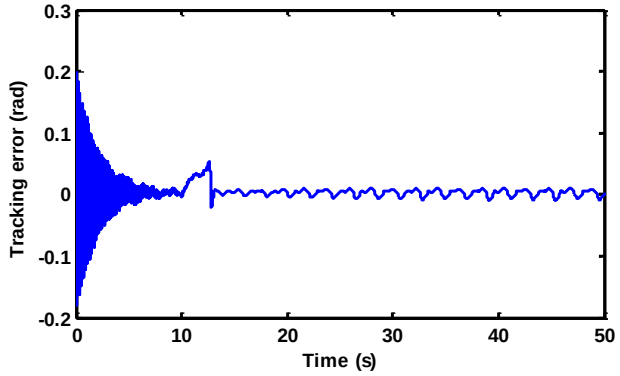

Figure 6.5. Pendulum 1 tracking error with accommodation

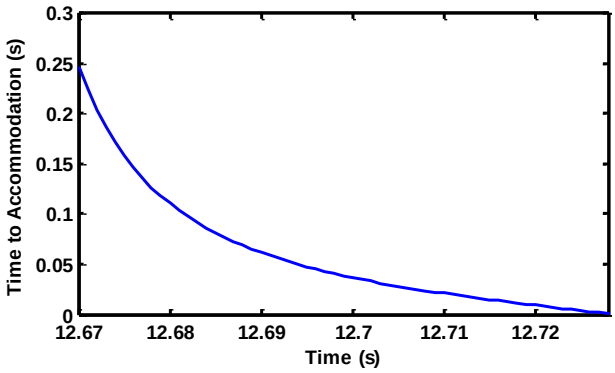

Figure 6.6. Estimated time-to-accommodation

7. CONCLUSIONS

The new decentralized fault detection and accommodation scheme presented in this paper is easy to implement on large-scale industrial systems, where significant amount of communication between subsystems due to state vectors is not possible or desirable. With the proposed scheme, the fault can be detected, its dynamics can be estimated by the detection OLA, and the accommodation can be performed to cancel the effects of the fault on all subsystems. Moreover, time-to-failure and time-to-accommodation are also approximated, which increases the system availability and reliability. These are all possible without the need for offline data from the system, known interconnection between subsystems, or the need for the overall measured or estimated state vectors. The only requirement of the proposed method is relatively accurate models of the subsystems consistent with other model-based schemes.

8. REFERENCES

[1] D. Miljkovic, "Fault detection methods: A literature survey," *MIPRO, 2011 Proceedings of the 34th International Convention*, pp. 750-755, 2011.

[2] H. Hammouri, M. Kinnaert, and E. H. El Yaagoubi, "Observer-based approach to fault detection and isolation for nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 44, pp. 1879-1884, 1999.

[3] R. J. Patton, J. Chen, and T. M. Siew, "Fault diagnosis in nonlinear dynamic systems via neural networks," *International Conference on Control '94.*, vol. 2, pp. 1346-1351, 1994.

[4] C. De Persis and A. Isidori, "A geometric approach to nonlinear fault detection and isolation," *IEEE Transactions on Automatic Control*, vol. 46, pp. 853-865, 2001.

[5] S. N. Huang and T. Kok Kiang, "Fault detection, isolation, and accommodation control in robotic systems," *IEEE Transactions on Automation Science and Engineering*, vol. 5, pp. 480-489, 2008.

[6] J. Bin, M. Staroswiecki, and V. Cocquempot, "Fault accommodation for nonlinear dynamic systems," *IEEE Transactions on Automatic Control*, vol. 51, pp. 1578-1583, 2006.

[7] N. Meskin and K. Khorasani, "Fault detection and isolation of distributed time-delay systems," *IEEE Transactions on Automatic Control*, vol. 54, pp. 2680-2685, 2009.

[8] Z. Yingwei, Z. Hong, S. J. Qin, and C. Tianyou, "Decentralized fault diagnosis of large-scale processes using multiblock Kernel partial least squares," *IEEE Transactions on Industrial Informatics*, vol. 6, pp. 3-10, 2010.

[9] L. Peng, Z. Bo, and L. Yuanchun, "Fault detection methods for reconfigurable manipulators via decentralized adaptive fuzzy nonlinear observer," *IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, pp. 1937-1941, 2010.

[10] M. A. Demetriou, K. Ito, and R. C. Smith, "Adaptive monitoring and accommodation of nonlinear actuator faults in positive real infinite dimensional systems," *IEEE Transactions on Automatic Control*, vol. 52, pp. 2332-2338, 2007.

[11] P. Panagi and M. M. Polycarpou, "Distributed fault accommodation for a class of interconnected nonlinear systems with partial communication," *IEEE Transactions on Automatic Control*, vol. 56, pp. 2962-2967, 2011.

[12] J. D. Boskovic and R. K. Mehra, "A decentralized scheme for accommodation of multiple simultaneous actuator failures," *Proceedings of the 2002 American Control Conference*, vol. 6, pp. 5098-5103, 2002.

[13] S. Jagannathan, *Neural network control of nonlinear discrete –time systems*. NY: CRC publications, 2006.

[14] H. Sunan, T. Kok Kiong, and L. Tong Heng, "Decentralized control design for large-scale systems with strong interconnections using neural networks," *IEEE Transactions on Automatic Control*, vol. 48, pp. 805-810, 2003.

[15] P. A. Ioannou, "Decentralized adaptive control of interconnected systems," *IEEE Transactions on Automatic Control*, vol. 31, pp. 291-298, 1986.

[16] Y. Qinmin, J. B. Vance, and S. Jagannathan, "Control of nonaffine nonlinear discrete-time systems using reinforcement-learning-based linearly parameterized neural networks," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 38, pp. 994-1001, 2008.

[17] H. Ferdowsi and S. Jagannathan, "A unified model-based fault diagnosis scheme for nonlinear discrete-time systems with additive and multiplicative faults," *50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, pp. 1570-1575, 2011.

[18] H. Ferdowsi, D. L. Raja, and S. Jagannathan, "A decentralized fault detection and prediction scheme for nonlinear interconnected continuous-time systems," *The 2012 International Joint Conference on Neural Networks (IJCNN)*, pp. 1-7, 2012.

[19] Y. Tang, M. Tomizuka, G. Guerrero, and G. Montemayor, "Decentralized robust control of mechanical systems," *IEEE Transactions on Automatic Control*, vol. 45, pp. 771-776, 2000.

V. Fault Diagnosis of a Class of Distributed Parameter Systems Modeled by Parabolic Partial Differential Equations

Hasan Ferdowsi and S. Jagannathan

Abstract

Many industrial systems are classified as distributed parameter systems (DPS) and the behavior of such systems is best described by partial differential equation (PDE) models. However, due to complex nature, a PDE model is traditionally transformed into a finite set of ordinary differential equations (ODE) prior to the design of control or fault detection schemes. As a result, significant approximations have to be made reducing the accuracy and reliability of the system. In this paper, the PDE representation of the system is directly utilized to construct a fault detection observer for DPS in contrast with the traditional fault detection observers which are based on the approximated ODE model of the system. A fault is detected by comparing the detection residual, which is the difference between measured and estimated outputs, with a predefined detection threshold. Once the fault is detected, an online approximator is activated to learn the fault function. An update law is introduced for updating the unknown parameters of the online approximator. The stability of the observer along with the online approximator is discussed analytically in the paper. Upon detecting a fault, the estimated fault parameters are compared with their failure thresholds to provide an estimate of the remaining useful life of the system. Further, a rigorous method for estimating the remaining useful life of the system in the presence of fault is introduced for DPS. The scheme is verified in simulations on a Lithium-ion battery system which is described by parabolic PDEs.

1. INTRODUCTION

Fault diagnosis has become an attractive research topic in the past couple of decades due to increased complexity of industrial systems and safety for such systems has become more important than ever. Among the different methods of fault diagnosis, model-based methods have become both popular and more suitable when a mathematical model of the system under consideration is either available or can be obtained [1] because they do not need extensive amounts of offline data and can operate online without requiring additional sensors. Therefore, model-based fault diagnosis schemes have been developed for lumped parameter systems based on their ordinary differential equation (ODE) representation using sliding mode observers [2], fuzzy observers [3], and adaptive observers [4].

A large number of industrial systems which involve heat transfer, fluid dynamics, electromagnetic, etc. are classified as distributed parameter systems (DPS). Application examples of such DPS include hydraulic systems, chemical processes, flexible robots, and aerospace systems. Due to the wide range of such systems and their important and sensitive role in the industry, reliable fault detection and diagnosis schemes are required to guarantee their safe operation.

The variables in DPS are defined over a continuous range of space [5], which makes them different from lumped parameter systems where each variable only evolves in time. The most comprehensive and accurate mathematical representation for these systems is given in terms of partial differential equations (PDEs). Limited work has been done on DPS when compared to the systems with ODE models, because dealing with PDEs is much more complicated due to boundary conditions and infinite number of states [5].

In order to simplify the problem of dealing with PDE model, it can be represented as an infinite set of ordinary differential equations (ODEs) and then apply the Galerkin's method to obtain an approximate finite dimensional ODE model [6, 7] which is then utilized for further analysis. Although this method has opened new doors to the problem of control and fault detection of DPS, it has a number of serious problems. First of all, these methods [6, 7] can possibly render inaccurate results since they are neglecting a significant portion of the system dynamics. Also there is no guarantee that the system

output can be generated as a function of only the states of the finite dimensional ODE system. Further, when a fault happens in the system, the PDE dynamics will change, which can make the approximated ODE model even more inaccurate.

Several fault detection methods for DPS have been proposed recently. For example, a learning systems approach is introduced in [8, 9] for fault detection of such systems whereas fault detection and isolation of such DPS with actuator faults are discussed in [10, 11], and a geometric approach is proposed for fault detection and isolation of dissipative parabolic PDEs [12]. However, all of them use the approach of transforming the PDE representation into an approximate finite dimensional ODE making these schemes [8-12] inaccurate and unreliable necessitating the need for a new technique for DPS.

In this paper, a fault diagnosis scheme by using the PDE representation of the DPS is introduced to increase the reliability and functionality of the entire system. In contrast with existing schemes [8-12], the PDE representation is not transformed into a finite dimensional ODE model before performing the fault detection. Instead, the detection observer is designed directly based on the PDE model. It is shown that the proposed observer can estimate both measured and unmeasured system states in the healthy operating conditions with a bounded error. Detection residual is generated by comparing the measured and estimated system outputs. Since the residual is bounded in healthy conditions, a fault can be detected by comparing it with a predefined detection threshold.

Upon detecting a fault, an online approximator, which is incorporated in the PDE observer, is activated to estimate the fault dynamics. An adaptive update law is proposed to tune the unknown parameters of the online approximator. The stability of both the proposed observer and the online approximator is investigated analytically. Furthermore, by using the parameter update law and comparing the estimated fault parameters with their failure thresholds an analytical formula for online estimation of the time to failure (TTF) or remaining useful life of the system is derived.

The effectiveness and stability of the proposed scheme is verified in a simulation example, a Lithium-ion battery system. The dynamics of the system is described by parabolic PDEs. The proposed PDE observer is utilized to estimate the unmeasurable

system state, Lithium concentration, and provide an estimation of the system output which is used to generate detection residual. Fault detection and approximation as well as remaining useful life estimation are successfully performed in simulations and the results are presented in the last section of this paper.

The paper is organized as follows: Section II describes the system and its model. Section III presents the PDE observer and Section IV discusses the online approximation of fault dynamics and Section V introduces the online failure prediction scheme. The verification of proposed scheme in simulations is presented in Section VI.

2. SYSTEM DESCRIPTION

Consider the class of nonlinear system described by the following normalized PDE

$$\frac{\partial x(z,t)}{\partial t} = a\frac{\partial^2 x(z,t)}{\partial z^2} + b(z)\frac{\partial x(z,t)}{\partial z} + c(z)x(z,t) \tag{1}$$

for $z \in (0,1)$and $t > 0$, with boundary conditions

$$\begin{cases} \dfrac{\partial x(0,t)}{\partial z} = qx(0,t) \\ \dfrac{\partial x(1,t)}{\partial z} = u(t) \end{cases} \tag{2}$$

where x is the state of the system, u is the control input applied at z=1, $a>0$ and q are constants, and $b(\cdot)$ and $c(\cdot)$ are smooth functions. Further, suppose that the only measurement is taken at the same end with actuation, i.e.

$$y(t) = x(1,t) \tag{3}$$

Now $b(z)$ can be eliminated from the equation by using the following transformation [13]

$$x(z,t) \to x(z,t)e^{-\frac{1}{2a}\int_0^z b(\tau)d\tau}$$

Next consider a fault in the system which can be modeled by $h(y(t),z,t)$. Then the system representation in (1) and (2) can be rewritten in the presence of a fault as

$$\frac{\partial x(z,t)}{\partial t} = a\frac{\partial^2 x(z,t)}{\partial z^2} + c(z)x(z,t) + h(y(t),z,t) \tag{4}$$

subject to the boundary conditions

$$\begin{cases} \dfrac{\partial x(0,t)}{\partial z} = qx(0,t) \\ \dfrac{\partial x(1,t)}{\partial z} = u(t) \end{cases} \tag{5}$$

Moreover, the fault function can be represented by

$$h\big(y(t),z,t\big) = \Omega(t-t_0)\bar{h}\big(y(t),z\big)$$

where $\Omega(t-t_0)$ is the time profile of the fault defined by

$$\Omega(\tau) = \begin{cases} 0 & , \ if\ \tau < 0 \\ 1-e^{-\kappa\tau} & , \ if\ \tau \geq 0 \end{cases}$$

where κ is an unknown constant determined by the growth rate of the fault. Although, this time profile is basically used to model incipient faults, it can also address abrupt faults by large values of κ.

The following standard assumption is needed in order to proceed.

Assumption 1: The fault function can be expressed as linear in the unknown parameters (LIP) [14], i.e. $h = w\sigma(y,z)$ where w is the vector of unknown parameters and $\sigma(\cdot)$ is a known nonlinear function which is bounded by $\|\sigma\| \leq \sigma_{max}$.

Next the PDE observer is designed to monitor the system states and output.

3. FAULT DETECTION OBSERVER

In order to detect a fault, an observer is utilized to estimate the system output in healthy conditions. Then the estimated and measured outputs will be compared to generate fault detection residual. Traditionally when dealing with PDEs, the system representation is transformed into an infinite set of ordinary differential equations, and then divided into an infinite dimensional fast and stable subsystem and a finite dimensional slow subsystem [8-12]. This way, the infinite dimensional part of the system is ignored and an ordinary ODE observer is designed to estimate the states of the finite dimensional part. Although this method has provided a basic solution to the problem of control and fault detection of DPS, it has several issues.

One of the most obvious shortcomings of the traditional method is the fact that a large part of the system dynamics has to be neglected, which can lead to inaccurate or even unreliable results. The other important issue arises from the limited number of measurements in a DPS which has infinite number of states. There is no guarantee that all the states of the slow subsystem can be measured in the actual system and an output observer might be useless, since the output of the system cannot always be represented as a function of slow subsystem states.

On the other hand, the ODE observer, designed based on the finite dimensional part of the system dynamics, is only reliable when the dynamics is completely known and fixed. In other words, unknown changes in the system dynamics, such as fault, can modify the eigenvalues, thus require a different transformation which will result in a different ODE representation. These issues motivated us to design a fault detection observer, directly based on the PDE model of the system.

Using the original PDE representation of the system and based on the Luenberger observer design, the following distributed parameter observer is proposed

$$\frac{\partial \hat{x}(z,t)}{\partial t} = a\frac{\partial^2 \hat{x}(z,t)}{\partial z^2} + c(z)\hat{x}(z,t) + L_1(z)\left(y(t) - \hat{y}(t)\right) \tag{6}$$

$$\begin{cases} \dfrac{\partial \hat{x}(0,t)}{\partial z} = q\hat{x}(0,t) \\ \dfrac{\partial \hat{x}(1,t)}{\partial z} = u(t) - L_2\left(y(t) - \hat{y}(t)\right) \end{cases} \tag{7}$$

$$\hat{y}(t) = \hat{x}(1,t) \tag{8}$$

where $\hat{x}$ and $\hat{y}$ are the estimated state vector and estimated output vector respectively, L_1 and L_2 are the output injection matrices of appropriate dimension which are used along with the output error, $y - \hat{y}$, to correct the observer error due to different initial conditions.

The next important issue is how to implement this fault detection observer in practice, because the online estimation is required while the PDE needs to be solved backward in time. In order to tackle this problem, the PDE observer will be discretized for implementation. Suppose the measurements are taken with a sampling rate of T, which means the output $y(t)$ is only available at times $t = kT$ for $= 1,2, \ldots$. Then the solution to the set of partial differential equations (6-8) is calculated in the time interval

(*t*-*T*,*t*) where t is the current time instant. For this purpose, y is assumed to remain constant in each sampling interval (*t*-*T*,*t*) and the final values of estimated states from the previous step, i.e. $\hat{x}(z, t-T)$ for $0 \le z \le 1$, are used as the initial values for solving the PDE in the time interval (*t*-*T*,*t*).

Since the only available measurement is *y*(*t*), the detection residual is defined as the difference between the measured and estimated outputs, i.e. $e = y - \hat{y}$. However, in order to analyze the stability of the observer, a state residual is also defined as $\tilde{x} = x - \hat{x}$. Then the residual dynamics in healthy operating conditions are described by

$$\frac{\partial \tilde{x}(z,t)}{\partial t} = a\frac{\partial^2 \tilde{x}(z,t)}{\partial z^2} + c(z)\tilde{x}(z,t) - L_1(z)\tilde{x}(1,t) \tag{9}$$

$$\begin{cases} \dfrac{\partial \tilde{x}(0,t)}{\partial z} = q\tilde{x}(0,t) \\ \dfrac{\partial \tilde{x}(1,t)}{\partial z} = L_2\tilde{x}(1,t) \end{cases} \tag{10}$$

As long as the system works in healthy operating condition, the residual dynamics should be stable and the state residual as well as detection residual must remain bounded. For this purpose we will look for a transformation [13] in the following form

$$\tilde{x}(z,t) = \xi(z,t) - \int_z^1 L(z,\tau)\xi(\tau,t)\,d\tau \tag{11}$$

which can transform the PDE in (9),(10) to

$$\frac{\partial \xi(z,t)}{\partial t} = a\frac{\partial^2 \xi(z,t)}{\partial z^2} - \kappa\xi(z,t) \tag{12}$$

$$\begin{cases} \dfrac{\partial \xi(0,t)}{\partial z} = 0 \\ \dfrac{\partial \xi(1,t)}{\partial z} = 0 \end{cases} \tag{13}$$

where κ is a user defined parameter introduced in order to keep the PDE system described in (12) and (13) stable. Therefore, we first substitute (11) in the system (9) and (10) and use Leibniz integral rule to get

$$\frac{\partial \xi(z,t)}{\partial t} - \int_z^1 L(z,\tau)\left[a\frac{\partial^2 \xi(\tau,t)}{\partial \tau^2} - \kappa\xi(\tau,t)\right]d\tau$$
$$= a\left[\frac{\partial^2 \xi(z,t)}{\partial z^2} + 2\frac{\partial L(z,z)}{\partial z}\xi(z,t) + L(z,z)\frac{\partial \xi(z,t)}{\partial z} - \int_z^1 \frac{\partial^2 L(z,\tau)}{\partial z^2}\xi(\tau,t)d\tau\right]$$
$$+ c(z)\left[\xi(z,t) - \int_z^1 L(z,\tau)\xi(\tau,t)d\tau\right] - L_1(z)\left[\xi(1,t)\right]$$

$$\begin{cases} \dfrac{\partial \xi(0,t)}{\partial z} + L(0,0)\xi(0,t) - \int_0^1 \dfrac{\partial L(0,\tau)}{\partial z}\xi(\tau,t)d\tau = q\left[\xi(0,t) - \int_0^1 L(0,\tau)\xi(\tau,t)d\tau\right] \\ \dfrac{\partial \xi(1,t)}{\partial z} + L(1,1)\xi(1,t) = L_2\left[\xi(1,t)\right] \end{cases}$$

By using integration by parts and rearranging the terms in the above equations, we arrive at

$$\frac{\partial \xi(z,t)}{\partial t} = a\frac{\partial^2 \xi(z,t)}{\partial z^2} + \left[3a\frac{\partial L(z,z)}{\partial z} + c(z)\right]\xi(z,t) - \left[L_1(z) + a\frac{\partial L(z,1)}{\partial \tau}\right]\xi(1,t)$$
$$+ \int_z^1 a\left[\frac{\partial^2 L(z,\tau)}{\partial \tau^2} - \frac{\partial^2 L(z,\tau)}{\partial z^2}\right]\xi(\tau,t)d\tau - \int_z^1 \left(c(z)+\kappa\right)L(z,\tau)\xi(\tau,t)d\tau \tag{14}$$

subject to

$$\begin{cases} \dfrac{\partial \xi(0,t)}{\partial z} = \left(q - L(0,0)\right)\xi(0,t) + \int_0^1 \left[\dfrac{\partial L(0,\tau)}{\partial z} - qL(0,\tau)\right]\xi(\tau,t)d\tau \\ \dfrac{\partial \xi(1,t)}{\partial z} = \left(L_2 - L(1,1)\right)\xi(1,t) \end{cases} \tag{15}$$

In order for the transformed PDE to be equivalent to the PDE described in (12) and (13), the following conditions which are obtained by comparing (14),(15) with (12),(13), must be satisfied

$$a\frac{\partial^2 L(z,\tau)}{\partial \tau^2} - a\frac{\partial^2 L(z,\tau)}{\partial z^2} = \left(c(z)+\kappa\right)L(z,\tau) \tag{16}$$

$$\begin{cases} \dfrac{\partial L(0,\tau)}{\partial z} = qL(0,\tau) \\ L(z,z) = q - \dfrac{1}{3a}\int_0^z \left(c(s)+\kappa\right)ds \end{cases} \tag{17}$$

$$L_1(z) = -a\frac{\partial L(z,1)}{\partial \tau} \tag{18}$$

$$L_2 = L(1,1) \tag{19}$$

Note that (16) describes a partial differential equation which has a unique solution [13] with the boundary condition defined in (17). Since (16) and (17) are time independent, its solution can be found offline and can be used to determine the observer parameters from conditions (18) and (19).

The next theorem discusses the stability of the residual dynamics in (9) and (10) under healthy operating conditions of the system.

Theorem 1 (PDE Observer Performance in the Healthy Conditions): Let the PDE observer introduced in (6) and (7) be used to estimate the states of the system described by (1) and (2), with L_1 and L_2 defined in (18) and (19). Then the state residual $\tilde{x}$ is exponentially stable (in the healthy operating conditions), if κ is selected to be positive.

Proof: It is already shown that there exists a unique transformation which can convert the residual dynamics in (9) and (10) into the target system of (12) and (13), if L_1 and L_2 are defined using (18) and (19). Since the transformation (11) is invertible [15], it only remains to be shown that the target system of (12) and (13) is asymptotically stable.

To investigate the stability of the PDE in (12) with boundary conditions in (13), the following Lyapunov function candidate is selected

$$V(t) = \frac{1}{2}\int_0^1 \xi^2(z,t)dz + \frac{1}{2}\int_0^1 \left(\frac{\partial \xi(z,t)}{\partial z}\right)^2 dz$$

Obviously V is positive definite, since $V(t) > 0$ for $t \geq 0$ if $\xi(z,t) \neq 0$ *or* $\frac{\partial \xi(z,t)}{\partial z} \neq 0$. Now we take the derivative of the Lyapunov function with respect to time and use integration by parts to get

$$\begin{aligned}\dot{V}(t) &= \int_0^1 \xi(z,t)\frac{\partial \xi(z,t)}{\partial t}dz + \int_0^1 \frac{\partial \xi(z,t)}{\partial z}\frac{\partial^2 \xi(z,t)}{\partial z \partial t}dz \\ &= \int_0^1 \xi(z,t)\frac{\partial \xi(z,t)}{\partial t}dz + \frac{\partial \xi(1,t)}{\partial z}\frac{\partial \xi(1,t)}{\partial t} - \frac{\partial \xi(0,t)}{\partial z}\frac{\partial \xi(0,t)}{\partial t} - \int_0^1 \frac{\partial^2 \xi(z,t)}{\partial z^2}\frac{\partial \xi(z,t)}{\partial t}dz\end{aligned}$$

By substituting (12) and (13) in the above equation and then using integration by parts again, we arrive at

$$\dot{V}(t) = -\kappa\int_0^1 \xi^2(z,t)dz - a\int_0^1\left(\frac{\partial^2\xi(z,t)}{\partial z^2}\right)^2 dz + (a+\kappa)\int_0^1 \frac{\partial^2\xi(z,t)}{\partial z^2}\xi(z,t)dz$$

$$= -\kappa\int_0^1 \xi^2(z,t)dz - a\int_0^1\left(\frac{\partial^2\xi(z,t)}{\partial z^2}\right)^2 dz - (a+\kappa)\int_0^1\left(\frac{\partial\xi(z,t)}{\partial z}\right)^2 dz$$

By applying the Poincare inequality [16] on the second term in the derivative of Lyapunov function, it can be rewritten as

$$\dot{V}(t) \le -\kappa\int_0^1 \xi^2(z,t)dz - \frac{a\pi^2}{4}\int_0^1\left(\frac{\partial\xi(z,t)}{\partial z} - \frac{\partial\xi(1,t)}{\partial z}\right)^2 dz - (a+\kappa)\int_0^1\left(\frac{\partial\xi(z,t)}{\partial z}\right)^2 dz$$

$$= -\kappa\int_0^1 \xi^2(z,t)dz - \left((\frac{4+\pi^2}{4})a+\kappa\right)\int_0^1\left(\frac{\partial\xi(z,t)}{\partial z}\right)^2 dz$$

The first derivative $\dot{V}$ is negative definite if κ is selected to be positive. Therefore the target system in (12) and (13) is asymptotically stable, which completes the proof.

Based on Theorem 1, fault detection can be performed by comparing the detection residual with a pre-defined threshold. This threshold needs to be determined based on the initial conditions of the system and the observer. When a fault occurs, the dynamics of the actual system will be changed, but the observer still estimates the system states and output based on the nominal system representation. Thus, the difference between measured and estimated states increases as the magnitude of fault grows. Once the detection residual reaches the detection threshold, a fault is declared active.

The next step in fault diagnosis is to determine the behavior of fault or approximating its dynamics, which allows further analysis of fault as well as estimation of remaining useful life of the system. To this end, an online approximator is added to the observer, which is discussed in the next section.

4. ONLINE FAULT APPROXIMATION

As mentioned above, the proposed detection observer is able to estimate the distributed system states and output with an asymptotically decreasing error in healthy operating conditions. When a fault occurs, the residual is no longer bounded and the fault is detected when the residual exceeds detection threshold. Upon detecting a fault, an online approximator is activated in the observer to estimate the fault dynamics. The online approximator is incorporated in the observer as

$$\frac{\partial \hat{x}(z,t)}{\partial t} = a\frac{\partial^2 \hat{x}(z,t)}{\partial z^2} + c(z)\hat{x}(z,t) + L_1(z)\left(y(t) - \hat{y}(t)\right) + \hat{h}(y(t), z; \hat{w}(t)) \tag{20}$$

where $\hat{h}$ is the output of online approximator defined by

$$\hat{h}(y(t), z; \hat{w}(t)) = \hat{w}^T(t)\sigma(y(t), z) \tag{21}$$

with $\hat{w}$ being the estimated matrix of unknown parameters.

By comparing the observer dynamics in (20) with the actual system dynamics in (4) and utilizing the Assumption 1, the residual dynamics after the detection of a fault can be represented by

$$\frac{\partial \tilde{x}(z,t)}{\partial t} = a\frac{\partial^2 \tilde{x}(z,t)}{\partial z^2} + c(z)\tilde{x}(z,t) - L_1\tilde{x}(1,t) + \tilde{w}^T(t)\sigma(y(t), z) \tag{22}$$

where $\tilde{w} = w - \hat{w}$ is the parameter estimation error. The stability of the proposed observer with the incorporated online approximator is discussed in the following theorem.

Theorem 2 (Fault Diagnosis Observer Performance): Let the proposed PDE observer in (20) with boundary conditions defined by (7) be used to monitor the system in (4) and (5), with the online approximator turned on upon detection of fault. If the parameter update law is defined as

$$\dot{\hat{w}}(t) = \beta\sigma(y(t),1)e(t) - \gamma\hat{w}(t) \tag{23}$$

where $\beta > 0$ is the learning rate and $\gamma > 0$ is the stabilizing term, then the FD residual, e, and the parameter estimation errors, $\tilde{w}$ are uniformly ultimately bounded (UUB), if the design parameters are selected such that

$$\kappa > \frac{\pi^2}{4} \quad and \quad \gamma > \beta\sigma_{\max}^2$$

Proof: First apply the transformation (11) on the residual dynamics

$$\frac{\partial \xi(z,t)}{\partial t} = a\frac{\partial^2 \xi(z,t)}{\partial z^2} - \kappa\xi(z,t) + \tilde{w}^T(t)\sigma(y(t), z) \tag{24}$$

$$\begin{cases} \dfrac{\partial \xi(0,t)}{\partial z} = 0 \\ \dfrac{\partial \xi(1,t)}{\partial z} = 0 \end{cases} \tag{25}$$

Consider the following Lyapunov function candidate

$$V(t)=\frac{1}{2}\int_0^1\xi^2(z,t)dz+\frac{1}{2}\int_0^1\left(\frac{\partial\xi(z,t)}{\partial z}\right)^2dz+\frac{1}{2\beta}\tilde{w}^T(t)\tilde{w}(t)+\frac{1}{2}\xi^2(1,t)$$

The derivative of this Lyapunov function is given by

$$\dot{V}(t)=\int_0^1\xi(z,t)\frac{\partial\xi(z,t)}{\partial t}dz+\int_0^1\frac{\partial\xi(z,t)}{\partial z}\frac{\partial^2\xi(z,t)}{\partial z\partial t}dz+\frac{1}{\beta}\tilde{w}^T(t)\dot{\tilde{w}}(t)+\xi(1,t)\frac{\partial\xi(1,t)}{\partial t}$$

Then we apply integration by parts and use the boundary conditions in (25) to get

$$\dot{V}(t)=\int_0^1\xi(z,t)\frac{\partial\xi(z,t)}{\partial t}dz-\int_0^1\frac{\partial^2\xi(z,t)}{\partial z^2}\frac{\partial\xi(z,t)}{\partial t}dz+\frac{1}{\beta}\tilde{w}^T(t)\dot{\tilde{w}}(t)+\xi(1,t)\frac{\partial\xi(1,t)}{\partial t}$$

The next step is to substitute (24) in the derivative of Lyapunov function and use integration by parts again to obtain

$$\begin{aligned}\dot{V}(t)=&-a\int_0^1\left(\frac{\partial^2\xi(z,t)}{\partial z^2}\right)^2dz-\kappa\int_0^1\xi^2(z,t)dz-(a+\kappa)\int_0^1\left(\frac{\partial\xi(z,t)}{\partial z}\right)^2dz\\&+\int_0^1\tilde{w}^T(t)\sigma(y(t),z)\left(\xi(z,t)-\frac{\partial^2\xi(z,t)}{\partial z^2}\right)dz-\kappa\xi^2(1,t)\\&+\frac{1}{\beta}\tilde{w}^T(t)\dot{\tilde{w}}(t)+\tilde{w}^T(t)\sigma(y(t),1)\xi(1,t)\end{aligned}$$

Now we employ the inequality $AB\le\frac{AA^T+B^TB}{2}$ where $A\in\mathbb{R}^{1\times n}$ and $B\in\mathbb{R}^{n\times 1}$

$$\begin{aligned}\dot{V}(t)\le&-(a-1)\int_0^1\left(\frac{\partial^2\xi(z,t)}{\partial z^2}\right)^2dz-(\kappa-1)\int_0^1\xi^2(z,t)dz-(a+\kappa)\int_0^1\left(\frac{\partial\xi(z,t)}{\partial z}\right)^2dz\\&-\kappa\xi^2(1,t)+\frac{\sigma_{\max}^2}{2}\tilde{w}^T(t)\tilde{w}(t)+\tilde{w}^T(t)\sigma(y(t),1)\xi(1,t)-\frac{1}{\beta}\tilde{w}^T(t)\dot{\tilde{w}}(t)\end{aligned}$$

To cancel out the positive term $\tilde{w}^T(t)\sigma(y(t),1)\xi(1,t)$, the parameter update law is selected as

$$\dot{\hat{w}}(t)=\beta\sigma(y(t),1)\xi(1,t)-\gamma\hat{w}(t)$$

In order to represent this update law in terms of available measurements of the system, the inverse of transformation (11) needs to be found. But an easier way is to set z=1 in (11) to get

$$\tilde{x}(1,t)=\xi(1,t)-\int_1^1L(1,\tau)\xi(\tau,t)d\tau=\xi(1,t)$$

Therefore, $e(t)=\tilde{x}(1,t)$ can be used instead of $\xi(1,t)$ in the update law to get (23). By substituting the proposed parameter update law in the derivative of Lyapunov function, it can be rewritten as

$$\dot{V}(t)\le -(a-1)\int_0^1\left(\frac{\partial^2\xi(z,t)}{\partial z^2}\right)^2 dz-(\kappa-1)\int_0^1\xi^2(z,t)dz-(a+\kappa)\int_0^1\left(\frac{\partial\xi(z,t)}{\partial z}\right)^2 dz-\kappa\xi^2(1,t)$$
$$+\frac{\sigma_{\max}^2}{2}\tilde{w}^T(t)\tilde{w}(t)-\frac{\gamma}{\beta}\tilde{w}^T(t)\tilde{w}(t)+\frac{\gamma}{\beta}\tilde{w}^T(t)w$$
$$\le -(a-1)\int_0^1\left(\frac{\partial^2\xi(z,t)}{\partial z^2}\right)^2 dz-(\kappa-1)\int_0^1\xi^2(z,t)dz-(a+\kappa)\int_0^1\left(\frac{\partial\xi(z,t)}{\partial z}\right)^2 dz$$
$$-\kappa\xi^2(1,t)-\frac{\gamma}{\beta}\tilde{w}^T(t)\tilde{w}(t)+\frac{\sigma_{\max}^2}{2}\tilde{w}^T(t)\tilde{w}(t)+\frac{\gamma}{2\beta}w^Tw+\frac{\gamma}{2\beta}\tilde{w}^T(t)\tilde{w}(t)$$

Finally, we apply the Poincare inequality and rearrange the terms in the above inequality to get

$$\dot{V}(t)\le -\left(\frac{4a+4\kappa+(a-1)\pi^2}{4}\right)\int_0^1\left(\frac{\partial\xi(z,t)}{\partial z}\right)^2 dz-(\kappa-1)\int_0^1\xi^2(z,t)dz$$
$$-\kappa\xi^2(1,t)-\frac{\gamma-\beta\sigma_{\max}^2}{2\beta}\tilde{w}^T(t)\tilde{w}(t)+\frac{\gamma w_{\max}^2}{2\beta}$$

Provided the conditions in Theorem 2 are satisfied, the derivative of Lyapunov function will be less than zero if the following are satisfied

$$\begin{cases}\left\|\dfrac{\partial\xi}{\partial z}\right\|^2>\dfrac{2\gamma w_{\max}^2}{\beta\left(4a+4\kappa+(a-1)\pi^2\right)},\ or\\ \|\xi\|^2>\dfrac{\gamma w_{\max}^2}{2\beta(\kappa-1)},\ or\\ \|e\|^2>\dfrac{\gamma w_{\max}^2}{2\beta\kappa},\ or\\ \|\tilde{w}\|^2>\dfrac{\gamma w_{\max}^2}{\gamma-\beta\sigma_{\max}^2}\end{cases}\qquad(26)$$

Therefore, the state estimation error and parameter estimation errors are UUB with the bounds defined by (26).

It has been shown so far, that the proposed method can perform fault detection and approximation with limited and adjustable error bounds. Using these results, we can move on to the next step which is the prediction of failure.

5. FAILURE PREDICTION

Despite the words fault and failure generally imply similar concepts they have different meanings in the field of fault diagnostics and prognostics. This difference originates from the fact that a fault does not necessarily make a system inoperable immediately after its occurrence and the system can continue to work with the existence of that fault for a limited time before it reaches a point called failure, when the system is no longer safe to operate or does not satisfy the minimum performance requirements. Therefore, once a fault is detected, it is crucial to determine the amount of time left before the system reaches a failure. This is generally called remaining useful life or time-to-failure prediction.

Another advantage of online model-based method of fault detection can be highlighted here, because the online estimation of fault and analytical parameter update law allows online estimation of failure time. For this purpose, the parameter update law in (23) is utilized along with failure thresholds for each of the fault parameters to derive a rigorous formula for time-to-failure estimation, which is discussed in the following theorem.

Theorem 3 (Time-to-Failure Prediction): Upon detection of fault through the PDE observer, the time-to-failure can be estimated using the following formula

$$TTF(t) = \min_i \left\{ \frac{1}{\gamma} Ln\left(\frac{\gamma \hat{w}_i(t) - \beta \sigma_i(y(t),1)e(t)}{\gamma \bar{w}_i - \beta \sigma_i(y(t),1)e(t)} \right) \right\} \tag{27}$$

Proof: In order to estimate the time of failure, the future values of the fault parameters should be estimated and then compared with the corresponding failure thresholds. Consider the parameter update law in (23) as a state space equation where the term $\beta\sigma(y(t),1)e(t)$ acts as an input. Then the solution to (23), which determines the value of estimated parameters in the future, can be described by

$$\hat{w}(t_f) = e^{-\gamma(t_f - t)}\hat{w}(t) + \int_t^{t_f} e^{-\gamma(t_f - \tau)}\beta\sigma(y(\tau),1)e(\tau)d\tau \quad for\ t_f > t$$

where t is the current time instant and t_f refers to future times. Since $\sigma(\cdot)$ is a bounded function and it has been shown that e remains bounded after the detection of fault, we will assume, as an approximation, that the term $\beta\sigma(y(\tau),1)e(\tau)$ remains constant. Now let $\hat{w}_i$ denote the i[th] element of the estimated parameter vector and suppose t_{f_i} is the time when the value of this element reaches its failure limit $\bar{w}_i$ for the first time. Then we will have

$$\hat{w}_i(t_{f_i}) = \bar{w}_i = e^{-\gamma(t_{f_i} - t)}\hat{w}_i(t) + \beta\sigma_i(y(t),1)e(t)\frac{1 - e^{-\gamma(t_{f_i} - t)}}{\gamma}$$

Define $TTF_i(t) = t_{f_i} - t$, which clearly refers to the time-to-failure corresponding to the i[th] parameter of the fault. By substituting this in the above equation and rearranging the terms we arrive at

$$\left(\hat{w}_i(t) - \frac{\beta}{\gamma}\sigma_i(y(t),1)e(t)\right)e^{-\gamma TTF_i(t)} = \bar{w}_i - \frac{\beta}{\gamma}\sigma_i(y(t),1)e(t)$$

$$\Rightarrow TTF_i(t) = \frac{1}{\gamma}Ln\frac{\gamma\hat{w}_i(t) - \beta\sigma_i(y(t),1)e(t)}{\gamma\bar{w}_i - \beta\sigma_i(y(t),1)e(t)}$$

This gives a formula for calculating time-to-failure for each fault parameter. Since the system fails if one of the fault parameters reaches its limit, the overall time-to-failure of the system is the small among time-to-failures of all parameters. This completes the proof.

6. SIMULATION RESULTS

To verify the proposed fault detection and prognosis scheme, it has been applied on a Lithium-ion battery in simulations using MATLAB. Two types of models are available for Lithium-ion batteries. One is the electrical circuit model which is described in terms of ordinary differential equations [17]. The other type of model is derived based on the chemical reactions inside the battery. This model is described in terms of partial differential equations and provides more information about the system can be found in [18]. Not only it is a more accurate model, but also it can provide estimates of Lithium concentration which is an unmeasurable distributed variable in the battery.

The system dynamics can be represented by the single particle model [19] which is in the form of following partial differential equations

$$\frac{\partial c_s^-}{\partial t}(r,t) = D_s^-\left(\frac{2}{r}\frac{\partial c_s^-}{\partial r}(r,t) + \frac{\partial^2 c_s^-}{\partial r^2}(r,t)\right) \tag{28}$$

subject to

$$\begin{cases} \dfrac{\partial c_s^-}{\partial r}(0,t) = 0 \\ \dfrac{\partial c_s^-}{\partial r}(R_s^-,t) = -\dfrac{I(t)}{D_s^- F a^- A L^-} \end{cases} \tag{29}$$

where c_s^- is the lithium concentration in anode, r is the radial coordinate, D_s^- is the diffusion coefficient in solid phase, R_s^- is the particle radius, F is the Faraday's constant, a^- is the specific interfacial surface area, A is the cell cross sectional area, L^- is the thickness of the anode, and I is the input current.

The battery model defined by (28) and (29) can be normalized using the following change in coordinates

$$r \to rR_s^- \qquad t \to t\frac{\left(R_s^-\right)^2}{D_s^-}$$

and then converted to the form of the PDE in (4) using the state transformation

$$c(r,t) = rc_s^-(r,t)$$

Therefore, the system dynamics in healthy conditions can be represented by

$$\frac{\partial c}{\partial t}(r,t) = \frac{\partial^2 c}{\partial r^2}(r,t) \tag{28}$$

subject to

$$\begin{cases} c(0,t) = 0 \\ \dfrac{\partial c}{\partial r}(1,t) - c(1,t) = -\dfrac{R_s^- I(t)}{D_s^- F a^- A L^-} \end{cases} \tag{29}$$

Note that the actual measurable output of the battery is the terminal voltage, but $c(1,t)$ can be calculated as a function of the terminal voltage. In order to simplify the notations, $c(1,t)$ is selected as the output. To seed a fault in the simulations, an internal fault is added to the PDE dynamics in (28) at the time $t = 100s$, resulting in

$$\frac{\partial c}{\partial t}(r,t)=\frac{\partial^2 c}{\partial r^2}(r,t)+0.1e^{-0.001(t-100)}c(1,t)r \qquad for\ t>100$$

The states of the system described by (28) and (29) are estimated using the proposed PDE observer which is described by

$$\frac{\partial \hat{c}}{\partial t}(r,t)=\frac{\partial^2 \hat{c}}{\partial r^2}(r,t)+L_1(r)\left(c(1,t)-\hat{c}(1,t)\right) \tag{28}$$

subjected to

$$\begin{cases}\hat{c}(0,t)=0\\ \dfrac{\partial \hat{c}}{\partial r}(1,t)-\hat{c}(1,t)=-\dfrac{R_s^- I(t)}{D_s^- Fa^- AL^-}-L_2\left(c(1,t)-\hat{c}(1,t)\right)\end{cases} \tag{29}$$

where L_1 and L_2 are obtained by numerically solving the PDE defined in (16) and (17). The observer parameters are selected as $a=1, \kappa=2$ which satisfy the stability condition in theorem 2. In order to implement the observer in simulations, it is solved in discrete time intervals with the length of one second and the observer is one second behind the actual system. In each time interval the observer takes $I(t)$ and $c(1,t)$ as inputs and is solved backward in time. Then the final values of that interval are used as initial values for the next time interval and so forth.

The drive cycle which is used in the simulations as the input to the system is depicted in Figure 6.1. The evolution of the distributed system variable, the Lithium concentration, is shown in Figure 6.2. As mentioned earlier, this variable is not measurable, but can be estimated by the PDE observer. It should be noted that estimated lithium concentration can be very useful in determination of the battery's state of charge. Although value of the system state at $r=0$ does not change due to the constant boundary condition, the lithium concentration at $r=1$ does not remain constant and its behavior significantly changes after the occurrence of fault.

The detection residual is generated by comparing actual and estimated value of the lithium concentration at $r=1$, i.e. $e(t)=c(1,t)-\hat{c}(1,t)$. It can be observed in Figure 6.3, that the detection residual remains bounded below the detection threshold as long as the system works in healthy conditions. Upon initiation of fault at $t=100s$, the residual starts to increase, because the estimated value of the system states will deviate from the actual values due to the change that the fault applies to the system dynamics.

Once the detection residual exceeds the detection threshold, a fault is detected and the parameter update law is immediately activated to learn the dynamics of the fault.

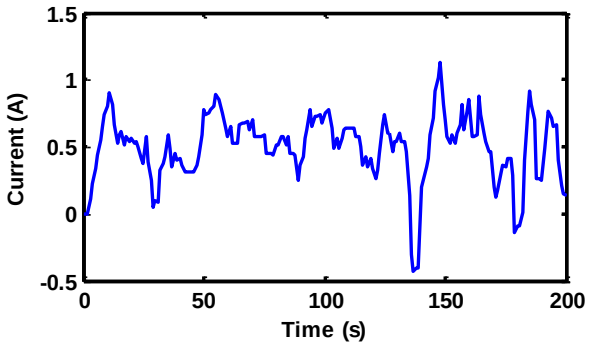

Figure 6.1. Input current

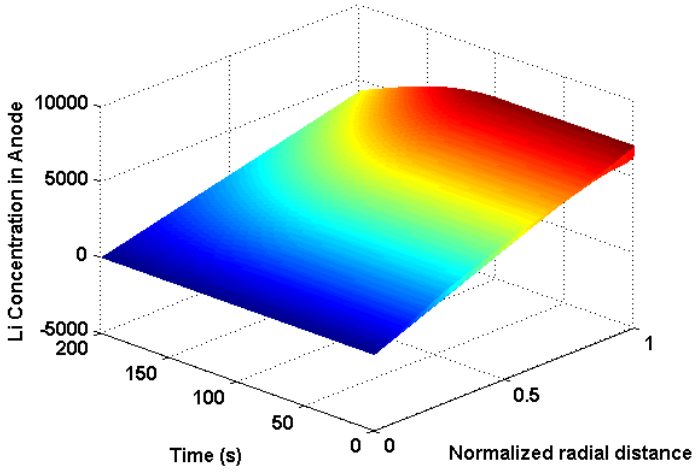

Figure 6.2. Lithium Concentration in Anode

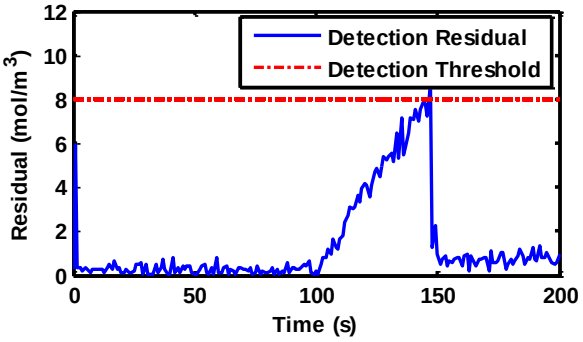

Figure 6.3. Detection residual and threshold

It can be seen in Figure 6.4, that the fault is seeded at time $t = 100s$. However, the estimated fault which is the output of the online approximator remains at zero until the detection happens, because the unknown parameters are initialized at zero and are not updated before the detection of fault. When the fault is detected at time $t = 147s$, the unpdate law starts to tune the unknown parameters and the estimated fault magnitude reaches the actual fault magnitude in less than 5 seconds. Accurate estimation of fault parameters not only results in accurate estimation of fault as seen in Figure 6.4, but also allows reasonable estimation of time to failure. Time to failure is estimated online using the formula given in (27) and the result is shown in Figure 6.5. Initial estimation of TTF is not accurate because the fault parameters have been initialized at zero. As the update law drives the parameters closer to their actual values, the estimation of time-to-failure becomes more accurate.

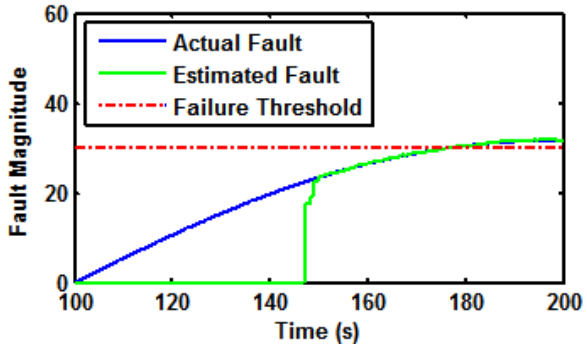

Figure 6.4. Actual and estimated fault

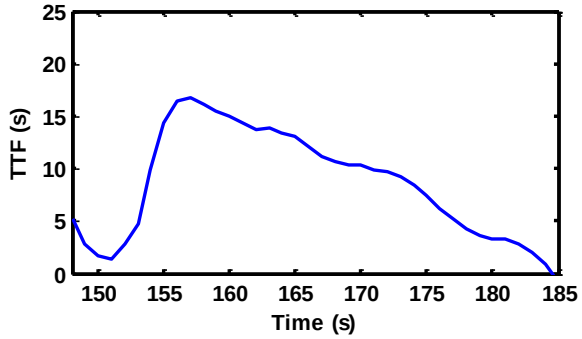

Figure 6.5. Estimated time-to-failure

7. CONCLUSIONS

Since the PDE model is not approximated with lower order models and it is directly used to construct the detection observer, the proposed scheme is more accurate in estimating the system states and more reliable in performing fault detection than the existing fault diagnostic methods for distributed parameter systems. As seen in the simulation example, it can also provide useful information about the unavailable system states. It was shown that if the stability conditions are satisfied with proper selection of design parameters, the observer will track the actual system states in healthy conditions and the adaptive update law will learn the unknown parameters of fault with a bounded error, which allows determination of time-to-failure. Accurate and fast fault detection enhances the reliability and decreases the maintenance costs while the time-to-failure determination increases the system availability.

8. REFERENCES

[1] J. J. Gertler, "Survey of model-based failure detection and isolation in complex plants," *IEEE Control Systems Magazine*, vol. 8, pp. 3-11, 1988.

[2] C. Edwards, S. K. Spurgeon, and R. J. Patton, "Sliding mode observers for fault detection and isolation," *Automatica*, vol. 36, pp. 541-553, 2000.

[3] R. J. Patton, J. Chen, and C. J. Lopez-Toribio, "Fuzzy observers for nonlinear dynamic systems fault diagnosis," *Proceedings of the 37th IEEE Conference on Decision and Control*, vol. 1, pp. 84-89, 1998.

[4] H. Ferdowsi and S. Jagannathan, "A unified model-based fault diagnosis scheme for non-linear discrete-time systems with additive and multiplicative faults," *Transactions of the Institute of Measurement and Control*, vol. 35, pp. 742-752, 2013.

[5] S. Omatu and J. H. Seinfeld, *Distributed parameter systems: theory and applications*. Oxford, Clarendon Press, 1989.

[6] P. D. Christofides, *Nonlinear and Robust Control of PDE Systems: Methods and Applications to transport reaction processes*. Boston, MA: Springer, 2001.

[7] A. Friedman, *Partial Differential Equations*. Dover Publications, Incorporated, 2008.

[8] M. A. Demetriou, "A model-based fault detection and diagnosis scheme for distributed parameter systems: A learning systems approach," *ESAIM: Control, Optimisation and Calculus of Variations*, vol. 7, pp. 43-67, 2002.

[9] A. Armaou and M. A. Demetriou, "Robust detection and accommodation of incipient component and actuator faults in nonlinear distributed processes," *AIChE Journal*, vol. 54, pp. 2651-2662, 2008.

[10] N. H. El-Farra and S. Ghantasala, "Actuator fault isolation and reconfiguration in transport-reaction processes," *AIChE Journal*, vol. 53, pp. 1518-1537, 2007.

[11] S. Ghantasala and N. H. El-Farra, "Robust actuator fault isolation and management in constrained uncertain parabolic PDE systems," *Automatica*, vol. 45, pp. 2368-2373, 2009.

[12] A. Baniamerian and K. Khorasani, "Fault detection and isolation of dissipative parabolic PDEs: Finite-dimensional geometric approach," *American Control Conference (ACC)*, pp. 5894-5899, 2012.

[13] A. Smyshlyaev and M. Krstic, *Adaptive control of parabolic PDEs*. Princeton University Press, 2010.

[14] S. Jagannathan, *Neural Network Control of Nonlinear Discrete –time Systems*. NY: CRC publications, 2006.

[15] A. Balogh and M. Krstic, "Infinite dimensional backstepping-style feedback transformations for a heat equation with an arbitrary level of instability," *European journal of control*, vol. 8, pp. 165-175, 2002.

[16] G. H. Hardy, J. E. Littlewood, and G. Polya, *Inequalities*. Cambridge university press, 1952.

[17] C. Min and G. A. Rincon-Mora, "Accurate electrical battery model capable of predicting runtime and I-V performance," *IEEE Transactions on Energy Conversion*, vol. 21, pp. 504-511, 2006.

[18] S. Santhanagopalan, Q. Guo, P. Ramadass, and R. E. White, "Review of models for predicting the cycling performance of lithium ion batteries," *Journal of Power Sources*, vol. 156, pp. 620-628, 2006.

[19] S. J. Moura, N. A. Chaturvedi, and M. Krstic, "PDE estimation techniques for advanced battery management systems — Part I: SOC estimation," *American Control Conference (ACC)*, pp. 559-565, 2012.

2. CONCLUSIONS AND FUTURE WORK

In this dissertation, adaptive nonlinear observers are designed to estimate system states and perform model-based fault diagnosis and prognosis. Novel parameter update laws guarantee the stability of the overall system during detection and prognosis and allow determination of time-to-failure upon detection of fault. Furthermore, a neural network (NN) approximator is used to develop an online outlier detection and removal scheme in order to prevent false fault alarms in the model-based schemes. One can use any online approximator in the diagnosis and prognosis schemes.

2.1. CONCLUSIONS

In the first paper, a fault diagnosis scheme is proposed that can handle both additive system faults and multiplicative actuator faults. The importance of this work lies in the fact that all industrial systems can be subjected to both types of faults. While the detection of fault regardless of its type is the crucial step, determination of fault type can narrow down the possible root causes of the detected fault and also allow the initiation of appropriate failure prediction scheme which will in turn improve system availability. The proposed scheme does not need large quantities of offline training data, it is generic, and can be applied to a wide range of systems provided a mathematical model is available. However, limited amount of input-output data is required to determine robust detection and identification thresholds. The only drawback of this scheme is that it requires all the system states to be available, which is not always possible. The fault diagnosis scheme can be applied on industrial systems with only a software upgrade and can help decreasing the repairing and down time costs.

In the second paper, a NN-based online outlier detection and removal scheme is presented and combined with a model-based fault detection scheme. While the existing methods of online outlier removal cannot provide a satisfactory performance in real applications because of the nonstationary environment, the proposed scheme can operate online in the presence of outliers, noise, fault, and change in operating conditions, since it does not require the system dynamics or underlying distribution of measurements to be fixed and known. The use of OIR scheme as a preprocessing unit guarantees removal of

outliers and so it can eliminate false fault alarms that are triggered by outliers. The estimation of outlier-free states is initially inaccurate due to random selection of neural network weights, but the learning algorithm quickly captures the system behavior and reduces the estimation error. Although in this paper, the OIR scheme is only combined with a model-based fault diagnosis scheme, it can also be used to enhance the performance of data-based fault detection.

The third paper proposes a decentralized fault prognosis scheme for spatially distributed systems. Unlike centralized schemes, asymptotic convergence cannot be guaranteed in the proposed decentralized scheme, but with a small sacrifice in accuracy a scheme is constructed which is more practical and easily implementable without the need for both complex centralized units and transmission of large amounts of data. Moreover, the scheme has higher reliability when compared to the centralized or consensus-based distributed schemes since the decentralized units are completely separated. The failure in one of the fault detector units does not interfere with the performance of other units which can prevent catastrophic failures in large networks. A single fault can be detected in all the subsystems that are significantly affected, and the origin of the fault is determined by the central fault isolation unit which only requires minimal information transmission.

The fourth paper presents a new decentralized fault detection and accommodation scheme which is easy to implement on large-scale industrial systems where significant amount of communication between subsystems is not possible or desirable. In the case of incipient faults, the accommodation allows uninterrupted operation of the system in the presence of fault for a limited time, which safely decreases the system downtime. By comparing estimated time-to-failure and time-to-accommodation, the appropriate decision to continue or stop the operation of system can be made. Therefore, the system repair can be consciously postponed to a suitable time by allowing the system to work in the presence of a fault without any risk of damage to system components or processes. These are all possible without the need for interconnection between subsystems to be known, or the need for the overall measured or estimated state vectors to be transmitted to all subsystems. The only requirement of the proposed method is the need for accurate models of the subsystems consistent with other model-based schemes.

In the last paper, the fault diagnosis of distributed parameter systems is investigated. The importance of this work is emphasized by the large number of systems that are characterized as distributed parameter systems, like the ones that involve hydraulics, electromagnetics, chemical reactions, etc. The major challenge in dealing with such systems is that they have unlimited number of states but limited number of sensors. The detection observer is designed without any approximations made in the system dynamics, thus making the estimation more accurate, decreasing the number of missed of false alarms in fault detection, while improving the accuracy in fault estimation and failure prediction. Additional sensors are not required since the observer can provide information about the unavailable system states which are useful for further analysis of the system and the root cause of the fault. Upon detection, adaptive estimation of fault parameters provides a reliable way to identify the fault which in turn helps in finding the location for early maintenance before the fault leads to further damage.

2.2. FUTURE WORK

The requirement for all the system states to be measurable, needs to be relaxed by using output observers. This is mainly to make our design more suitable for practical implementation, since many system states might not be measurable. Even if all the system states are measurable, it is desirable to perform the fault diagnosis using only the existing sensors. The future work regarding the decentralized fault accommodation scheme is related to the class of systems under consideration, as the current scheme can only handle systems in the form of controllable canonical form. Consequently, novel schemes have to be developed to cover broad class of affine systems.

Another part of the future work is the hardware implementation of the proposed decentralized fault diagnosis and prognosis scheme as well as the PDE-based fault diagnosis scheme. Although the effectiveness of these schemes was illustrated through the use of simulation examples, hardware implementation is a mandatory step for any industrial design and can determine possible problems or shortcomings of the scheme, thus helping us to improve it.

The PDE-based diagnosis scheme for distributed parameter systems is a new area and has a lot of room for improvement. In this dissertation, a scheme was proposed only

for the class of systems which are modeled by linear parabolic PDEs. The fault diagnosis of systems based on other types of PDEs, like hyperbolic and elliptic PDEs, is still an open problem which is definitely worth investigating. Also the observer design and fault detection of systems with nonlinear PDEs is a part of the future work. Another important problem to be targeted is the fault accommodation of the distributed parameter systems based on the current detection scheme.

VITA

Hasan Ferdowsi was born in Isfahan, Iran, in 1984. He earned his Bachelor of Science and Master of Science in Electrical Engineering from Isfahan University of Technology in 2007 and 2010, respectively. In August 2010 he joined Missouri University of Science and Technology as a PhD student. During his stay at the university, he worked under supervision of Professor Jagannathan Sarangapani and received Doctor of Philosophy degree in Electrical Engineering in December 2013.