NOTE TO USERS

Fault Diagnosis in Spacecraft Attitude Control System

Hamed Azarnoush

A thesis

in

The Department

of

Electrical and Computer Engineering

Present in Partial Fulfillment of Requirements
For the Degree of Master of Applied Science (Electrical Engineering) at
Concordia University
Montreal, Quebec, Canada

August 2005

Your file Votre référence
ISBN: 0-494-10232-2
Our file Notre référence
ISBN: 0-494-10232-2

Abstract

Fault Diagnosis in Spacecraft Attitude Control System

Hamed Azarnoush

This thesis presents a method for fault diagnosis in spacecraft attitude control system. The focus of this diagnosis problem will be on the actuator components. The considered actuator is assumed to be a reaction wheel. Since the attitude model of the spacecraft to be used is a coupled three-axis, three reaction wheels are required to provide actuation signals in this system. A high-fidelity model of the reaction wheel is considered which incorporates the effect of most disturbances involved in practice. A controller is designed to guarantee the stability of the signals in the system. This controller is designed according to nearly ideal model of the reaction wheel in a single axis system. The performance of this controller is validated for the three-axis nonlinear model of the spacecraft attitude control system using nonlinear model of the reaction wheel. The contribution of this thesis is designing a nonlinear observer which can diagnose the faults in the reaction wheel components. The nonlinear observer is designed based on a linear observer obtained through linearization of the reaction wheel model. This nonlinear observer serves as the fault diagnosis module which provides a model of the reaction wheel. This module generates two residual signals which determine when the system is under normal operating conditions and when the system is faulty. The residual signals can detect three fault types in three different components of the reaction wheel. One of these faults can be isolated from the other two as well. The behavior of the linear observer subject to the same anomalies and faults is also investigated to justify the development of the proposed nonlinear diagnosis module.

Acknowledgement

I am grateful to my academic advisor, Prof. K. Khorasani, for his support and guidance throughout my work. His inspiring advice and criticism guided this thesis all the way to the end. This work would have never been done without his great vision, experience and insight.

I also would like to thank all my friends who have helped, supported and made my time enjoyable in Concordia University.

I want to express my gratitude to the organizations that supported this work. These include the Natural Science and Engineering Research Council and the Canadian Space Agency.

Last, but not least, I am sincerely thankful to my family for their patience and incredible support. I am lucky to have a wonderful, loving and supportive family and I want to thank them all.

This thesis is dedicated to my family

Contents

List of Figures

List of Tables

Chapter 1

1 Introduction

1.1 Statement of the Problem

In this thesis, the problem of Fault Detection and Isolation (FDI) in a spacecraft attitude control system is addressed. The attitude control system (ACS) is composed of different components such as sensors (e.g. horizon sensors, sun sensors, magnetometers, etc.) and actuators (e.g. momentum wheels, electromagnets, etc.). The problem of FDI can be developed for different components of the ACS. The focus of fault diagnosis in this thesis will be on the actuators of the ACS. Reaction wheels play an important role in the attitude control system. They are, on the other hand, very susceptible to faults. This is the main motivating factor that the solutions to FDI problem in these devices are sought in this thesis. Therefore, the problem will be that of constructing a fault diagnosis module for the reaction wheel. This module will act as the supervising unit in the FDI scheme. Any time that a fault or anomaly takes place this module is responsible of generating a warning signal to report the existence of a fault in the system.

Since the model of the reaction wheel employed here is a high-fidelity model, this module should be designed in a way that it can sense minor faults in different parameters of this device. The ACS is also under the influence of various disturbances and noises.

The diagnosis scheme should not erroneously react to these disturbances or noises as if a fault has occurred in the system. Therefore, there is usually a trade-off between accurately detecting small faults or anomalies, and filtering the noises or disturbances in the system.

1.2 Justification

The problem of fault detection, isolation and recovery (FDIR) has been a subject of interest in recent years. Modern control systems are becoming more and more complicated and control algorithms more and more sophisticated. As a result the issues of availability, cost efficiency, reliability, operating safety and environmental protection are of major significance. For safety-critical systems, the consequences of faults can be extremely serious in terms of human mortality, environmental impact and economic loss. Therefore there is growing need for on-line supervision and fault diagnosis to increase the reliability of such safety-critical systems. Early indications confirming which faults are developing in the system can help avoid system breakdown, mission abortion and catastrophes. Space vehicles can be named as an example of safety-critical systems.

Advanced computers have allowed for more tasks to be accomplished onboard the space vehicles. Ground activities such as navigation and maneuver planning, command planning and sequencing, data summarization, and fault diagnosis and recovery can all be done autonomously onboard the spacecrafts.

The distance may be the most significant factor that makes the existence of an onboard autonomy essential. The farther the spacecraft is from the ground, the less knowledge is available about its present environment. Also the distance causes huge

delays in communication between the ground and the spacecraft. For deep space probes near planets like Pluto a radio signal may require five hours to accomplish its trip.

The importance of embedding fault diagnosis into goal oriented autonomous spacecrafts has become critical. Future generations of spacecrafts need to show proper reaction to unexpected events such as component failure or environmental interaction. Most currently used controllers react in different situations according to their coded routines. This is impractical when the spacecraft is facing an unexpected event. Embedded planners make it possible for the controller to reschedule the mission. This is sometimes called as the second generation diagnosis system, where the first generation is identified by the fuzzy if-then rules. Model-based diagnosis (MBD) is a powerful method of fault diagnosis. In MBD, a mathematical model of an aspect of the vehicle is generated and its behavior is compared with that of the real system being observed. Fault is detected according to observation of the difference between the model behavior and the real plant.

Onboard fault diagnosis can detect and isolate spacecraft faults in a very short period of time. This ability to fulfill these tasks onboard allows the spacecraft to reschedule the mission in presence of major anomalies and still maintain as many mission objectives as possible.

The probability of fault increases with the time needed to accomplish the mission. Hence there is growing need to have a health monitoring mechanism onboard. This need for a self-healing autonomy has been the motivation to generation of different methods in different fields including system identification, robust and adaptive control, and system health monitoring.

In remote operations it is valuable from different aspects to have a system that diagnoses and fixes unexpected problems. For operations that last for a long period of time, functions such as navigation and subsystem management are much desirable to be performed without intervention from a ground center. Some of the advantages are as follows:

- Sometimes there is a short period of time available to establish communication between the spacecraft and ground station, even under normal operation of the system without having any fault present,
- Reduced cost of ground-based support during the long mission,
- Eliminating the long round trip delays due to large distance between the spacecraft and the ground, and
- Eventual loss of communication with the ground due to the existing unpredictable environment around the space vehicle.

The main goal of an autonomous operation should be to maintain the spacecraft's safety and to perform the critical functions in priority.

1.3 Motivation for This Research

Reaction wheels play a vital role in spacecraft attitude control system. These wheels are components that are vulnerable to faults. Therefore, it is of utmost importance to develop a fault diagnosis algorithm for these sensitive components. Unfortunately there are not a lot of references in which this problem is addressed.

In spacecraft three reaction wheels are needed to control the attitude in all directions. However, four reaction wheels are usually mounted on the spacecraft. Three wheels actively participate in the attitude control and the fourth wheel is a redundant device to be

applied during the time that one of the wheels is experiencing a trouble. A comprehensive analysis on the behavior of the wheels and their sensitivity to faults, and reconfiguring or healing the system according to type of the fault, may reduce the need for the use of the fourth wheel as a redundant component.

In most of the literature the problem of FDI is discussed for actuators that have simple models. The fault in the actuators is usually considered as unexpected variations in the output signal of the actuator. In this thesis, faulty components are assumed to be in the actuator model components. In other words, the problem of fault diagnosis at the actuator level is discussed in more detail.

An observer-based method is applied as a model-based approach to solve the FDI problem. This work may also motivate other researchers in applying other approaches to fault diagnosis in this device.

1.4 Methodology

Development of fault detection, isolation and recovery (FDIR) algorithms has received a lot of attention in the past two decades. During this time, numerous FDIR approaches have been suggested by different researchers [1]-[51]. The traditional approach to fault diagnosis is based on hardware redundancy methods in which multiple redundant components are used to detect the fault when there is inconsistency between the outputs of different components. Another approach known as analytical redundancy is introduced to solve the problem of FDIR. In this approach analytical relationships between different variables of the system are used for the purpose of fault diagnosis. When these relationships are violated, it means that there is a faulty component in the system. Since analytical redundancy approaches use the mathematical model of the

system as the point of reference to diagnose the fault, these approaches are usually called model based approaches.

In this thesis a model-based approach is applied to fault diagnosis in the spacecraft attitude control system. There is a variety of different model-based approaches suggested by different researchers [8]-[17]. In this thesis an observer-based method is used as a model-based approach to FDI problem. In this method, a linear observer is designed as a conventional approach to solve the problem. The performance of this observer will be evaluated and analyzed. Since the attitude control system is a nonlinear model, the linear observer has some inefficiency in predicting the model behavior and consequently solving the FDI problem. Hence a nonlinear observer is designed as an alternative to compensate the weaknesses of the linear observer. Therefore, the performance of this linear observer is used as a benchmark to evaluate that of the nonlinear observer.

When an observer is applied to monitor the performance of a system, the first step is to determine which part of the system should be supervised. There may be different components in any system that can be vulnerable to fault. Before the observer is designed, it should be clarified that which components are fault-prone or which ones are going to be monitored. The focus of diagnosis may be one component, a set of components or the whole system.

In this thesis, the observer is designed to monitor the attitude control system (ACS). This system is composed of different components. The FDI goal might be diagnosing the fault in sensors, actuators or the spacecraft itself. It could also focus on a group of these components. In this thesis, the FDI problem is centered on the actuator of the ACS. Thus

the observers will be designed to inspect the behavior of this component in the spacecraft attitude control system.

In spacecraft, different devices may be used as actuators to control the attitude. Reaction wheels are devices that are widely used in spacecraft as the actuators in the ACS. These devices are vulnerable to faults. The model that is used for the reaction wheel is a high-fidelity model that incorporates the effect of most disturbances and details that are involved in practice. In this model it is assumed that three components are subject to fault. The capability of the designed observers in diagnosing these faults will be investigated.

Since the FDI problem is concerned with the attitude control system, to reach the final goal of diagnosing the faults in the system the first step is stabilization of the ACS. Thus, the first problem, to be discussed, deals with designing a controller for the ACS. The spacecraft attitude control system is a fully nonlinear three-axis system with coupling between different axes. Since it is too complicated to start designing a controller for this system, a simplified version of this system is applied to determine the stabilizing control law. In this prototype model, all nonlinearities are approximated by linear components. The three-axis model will be reduced to a single-axis model. This is accomplished by assuming that the spacecraft is symmetric. The reaction wheel model is also replaced by a nearly ideal linear counterpart. After finding the simplified model, a controller can be designed for this linear model according to classical control laws. The next step is to test the performance of this control law and verifying if it satisfies the desired system performance specifications. After validating the performance of this controller for the linear model, it is used as the control module in the original nonlinear model. By

experimentally adjusting the parameters of this controller for the nonlinear model, this controller can be optimized and then used satisfactorily in the fully nonlinear system.

1.5 Contribution of Thesis

In this thesis, a nonlinear observer is designed as the fault detection and isolation module in a high-fidelity model of a reaction wheel in spacecraft attitude control system. This module is applied to detect three types of faults in a nonlinear model of the reaction wheel, used as an actuator in the spacecraft attitude control system. Since these devices are very sensitive devices and they are very vulnerable to faults, a fault diagnosis algorithm to monitor these devices is on high demand. To justify the use of complexities involved in a nonlinear fault diagnosis module, the performance of the nonlinear observer is compared with that of a linear one. In this comparative analysis, the advantages of the nonlinear observer over the linear one are illustrated and emphasized.

Unfortunately not a lot of resources or previous published research on this specific subject exists in the literature. Hopefully, this thesis will open a new sequence of future work dealing with this problem.

1.6 Outline of Thesis

This thesis is organized as follows. Chapter 2 presents a comprehensive analysis of different approaches used in fault diagnosis. Chapter 3 explains the spacecraft attitude control system in which the fault diagnosis problem is defined. All component models are explained in this chapter. In Chapter 4, the theory and design of state estimators and the fault diagnosis architecture are described. Chapter 5 presents simulation results obtained

from applying the designed diagnostic scheme. Finally concluding remarks and future work are included in Chapter 6.

Chapter 2

2 Fault Detection, Isolation and Recovery (FDIR)

In this chapter, various approaches to Fault Detection, Isolation and Recovery (FDIR) are introduced. It starts by explaining the concept of fault tolerance and the role of FDI in fault tolerant control systems. This follows by expressing desirable attributes of a diagnostic system. Finally a comprehensive analysis of different approaches to FDI is presented.

2.1 The Concept of Fault Tolerant Control Systems

It is significantly important for an operating system to continue performing satisfactorily and fulfill specified functions when a fault exists in the system. However, when there might be performance degradation in the system it should be capable of maintaining the primary objectives as well as giving the operator sufficient time to find a cure for the problem. Since a system needs to possess desired reliability, performance in faulty conditions, maintainability and survivability, the fault tolerance concept has received increasing attention in the past few years.

Some characteristics of a fault tolerant control are as follows [1, 2]:

Generality: The fault tolerant mechanism should be able to handle a variety of fault sources and to handle both transient and permanent failures.

Appropriate response: the mechanism is supposed to respond to changes in the environment, mission phases and system states within an acceptable period of time.

Versatility: The fault tolerant mechanism must be able to consider various modes of adapting the system in response to anomalous behaviors. This adaptation for example can be done in one of the following ways: (i) switching from one fault-handling mode to another, (ii) modifying parameters of the currently used fault-handling mode, or (iii) modifying service attributes.

Recent failure history: The basic idea is that if the recent failure history, informs that the probability of fault in the near future is high, the fault tolerant mechanism should switch from an optimist fault handling mode to a pessimist one. The reverse switching should be executed when the probability is low.

In order to achieve the objectives of fabricating a fault-tolerant system, a powerful fault diagnosis scheme is required. To explain the role of FDI in fault tolerant control systems, first the FDI terminology is explained.

2.2 Fault Diagnosis Terminology

Although there is not a common terminology in all the literature, in this thesis we tried to use the one that is consistent with most text books and articles [3].

A fault is an unexpected change of system function although it may not represent physical failure or breakdown. Such a fault or malfunction hampers or disturbs the normal operation of an automatic system, thus causing an unacceptable deterioration of the performance of the system or even leading to dangerous situations. The term "fault"

is used rather than failure to denote a malfunction rather than catastrophe. The term failure suggests complete breakdown of a system component or function, whilst the term fault may be used to indicate that a malfunction may be tolerable at its present stage. A fault must be diagnosed as early as possible, even if it is tolerable at its early stage, to prevent any serious consequences.

A monitoring system which is used to detect faults and diagnose their location and significance in a system is called a "fault diagnosis system". Such a system normally consists of the following tasks:

Fault detection: to make a decision that either something has gone wrong or everything is fine.

Fault isolation: to determine the location of the fault, (e.g. which sensor or actuator has become faulty)

Fault identification: to estimate the size and type or nature of the fault.

After becoming familiar with the fault diagnosis concepts, the role of FDI in fault tolerant control systems and desirable characteristics of a diagnostic system can be discussed.

2.3 The Role of FDI in Fault Tolerant Control Systems

Fault diagnosis in a system helps detect and isolate the fault. After detection of fault and possibly locating the source of fault, the system can be reconfigured or restructured. In this case healthy and non-faulty components are chosen to take role in system operation. In some cases a pre-calculated controller is switched on or some parameters of the controller are changed to fit the new situation. The main problem in fault tolerant control is on-line reconfiguration or restructuring of the controller. Achieving this objective requires a lot of detailed information about changes in the system operating

point or different parameters caused by component fault or normal process changes. The task of FDI unit is to acquire this information for fault diagnosis and the task of the supervision unit is to choose the appropriate reconfiguration or restructuring strategies according to the diagnostic information it receives from the FDI unit. In real applications, the probability of occurrence of some faults is much smaller than that of the others. However, there is no guarantee that those types of faults do not occur in the system. Some approaches do not require all detailed information about the nature of faults, time and location of their occurrence for online reconfigurable controller design. However, the reliability of the system should be ensured by an FDI module. In these approaches the reconfigurable control is only switched on when the FDI unit detects a fault. Next, desirable attributes of a fault diagnosis system are described.

2.4 Desirable Attributes of a Diagnostic System

A desirable fault diagnosis system should meet some expectations. Some desirable attributes of a diagnostic system are explained in detail:

Early detection and diagnosis: It is important that faults in the system be diagnosed quickly and accurately enough. However this is conflicting with having tolerable performance during normal operation of the system.

Isolability: Isolability is the capability of the diagnostic system in distinguishing the source of fault. However there is a trade-off between having a high degree of isolability and rejecting modeling uncertainties. A classifier with a high degree of isolability may be weak in rejecting modeling uncertainty.

Robustness: There should be some robustness to noise and modeling uncertainty in the diagnostic system. Robustness precludes the deterministic isolability tests in which the thresholds are very close to zero.

Novelty Identifiability: The diagnostic system's ability to recognize novel malfunctions in the system is known as novelty identifiability. The novel malfunctions in the system should not be categorized as other known malfunctions or should not be treated as if the system is under normal operation conditions.

Multiple Fault Identifiability: This is a very difficult requirement for a classifier to be able to identify multiple faults. Due to nonlinearities and coupling interactions generally existing in spacecraft models, it is very difficult to model the combined effect of faults.

Explanation Facility: A diagnostic system should be able to explain where a fault originated and how it propagated in the system.

Adaptability: The operating conditions of the system change due to disturbances or environmental changes. A diagnostic system should be adaptable to changes in the system.

Reasonable storage and computational requirement: Less computationally complex algorithms and implementations lead to faster decision makings. However they necessitate high storage requirements. A reasonable compromise between these two requirements should be made.

2.5 A Review of Different Diagnostic Approaches

The analysis of various diagnostic techniques starts with comparing and contrasting different approaches from different aspects. First we start by categorizing the FDI approaches based on hardware or analytical redundancy.

2.5.1 Hardware versus Analytical Redundancy

One traditional approach to fault diagnosis is hardware redundancy. In this method, redundant components are utilized in the diagnosed system. These redundant components can be sensors, actuators, etc. By polling among these redundancies, any faulty component can be located through inconsistent information that is being generated. Although this approach has been widely applied, it has some major problems. Obviously using extra equipments costs more and creates maintenance problems as well. Furthermore, redundant devices need additional space to be accommodated.

Although benefiting redundant devices is a reliable way of detecting the fault, replicating hardware may not be the best idea for solving this problem. One way to inspire the idea of hardware redundancy without confronting its disadvantages is to use analytical redundancy. In this approach, an analytical replica of the model is produced. Since this model can be generated in the computer, the problems involving with a hardware component are avoided. This analytical model is built according to the relationships and rules that should hold between different variables in the system during normal operating conditions. Anytime these rules are breached in the system, the fault diagnosis scheme detects an inconsistency between the system variables and those of the analytical model and consequently detects the fault. Since in the analytical redundancy approach, a mathematical model is generated to inspect the behavior of the system, this approach is usually referred to as model-based approach. The idea of using analytical redundancy was first suggested by Beard [4]. Subsequently, the geometric interpretation of this approach was offered by Jones [5] and Massoumnia [6]. Other model-based approaches were developed by a number of investigators that will be introduced later.

Another way to classify the methods used in fault diagnosis is according to the form of process knowledge required. The form of process knowledge divides current methods into process model-based techniques and process history-based techniques [7].

2.5.2 Process-model-based Techniques

These methods mainly rely on causal or model-based knowledge. The source of knowledge in these methods is a deep understanding of the process. The process model-based techniques can be divided to quantitative methods and qualitative causal methods.

2.5.2.1 Qualitative Causal Methods

These methods rely on cause and effect reasoning about system behavior. Signed digraphs and fault trees are among these methods. In fault trees [8], a primary root of the fault is sought through backward chaining from the observed effects of the fault. In these methods a lot of hypotheses are generated and a lot of qualitative ambiguities are involved. This contributes to a more uncertain decision making. These methods are inspired from human brain analytical reasoning which makes them more straightforward.

2.5.2.2 Quantitative Methods

These methods rely on the mathematical relationship between different variables in the system. In these approaches a model gathers all the information about the system. These models generate a signal called residual. Residuals are usually a criterion to show divergence of the system variables from the mathematical model variables. Hence, they can be used for the purpose of fault diagnosis. Different residual-based approaches have been suggested by various researchers.

The observer-based method, for instance, was investigated in depth by Frank [9]. Most of the nonlinear observer-based approaches assume that the faults can be decoupled from modeling uncertainties and disturbances through a proper coordinate transformation (Krener and Isidori [10]). While other approaches assume that the effects of modeling uncertainties is less significant than that of faults. Since nonlinear systems are very complex, numerous directly designing approaches have been developed. For example Besancon and Hammouri [11] studied the observer design problem utilizing the solution of Riccati equation for Lipschitz nonlinear systems. High gain techniques have also been used as in Praly *et al* and Khalil [12-14].

The parity relation approach was originally proposed by Mirnovsky [15]. In this approach residual generator is generated through consistency checking on system input and output data over a time window. Due to limited availability, this approach was later independently proposed by Chow and Willsky [16]. Gertler [17] expressed this method in the discrete time domain.

2.5.3 Process-history-based Techniques

In contrast to the model-based approaches in which a prior knowledge of the system is required, process-history based approaches need a large amount of process history data. There are different ways in which process data can be presented to the system. This is known as feature extraction. The nature of this extraction can be either quantitative or qualitative. Next, different feature extraction methods will be introduced [7].

2.5.3.1 Qualitative Methods

Two major methods that apply qualitative history data are the expert systems and trend modeling methods.

Rule-based methods

These methods are the primary focus of expert systems. An expert system is basically capable of solving a problem in a very narrow domain. In these methods there is an explicit mapping of known symptoms to root causes.

The main advantages of developing expert systems in fault diagnosis are: ease of development, transparent reasoning, the ability to reason under uncertainties and the ability to present explanation for the provided solutions.

Several papers have discussed the application of expert systems in fault diagnosis of specific systems. Initial attempts applying these methods can be found in Henley [18], Chester, Lamb and Dhurjati [19] and Venkatasubramanian [20]. The objectives pursued by them were twofold. The system assigns the problem to either operator error, equipment failure or disturbance. Second, the expert system finds a solution to return the system back to normal operating conditions. Ideas on knowledge-based diagnostic systems based on the task framework can be found in Ramesh, Davis, and Schwenzer [21]. Basila, Stefanek, and Cinar [22] have developed a supervisory expert system that uses object based knowledge representation to represent heuristic and model-based knowledge.

Qualitative Trend Analysis

Another approach in qualitative feature extraction is the abstraction of trend information. Trend modeling can be applied to provide explanations for the occurrence of

different events in the system, diagnose malfunctions and predict future states. In this analysis there are two steps: (i) Identification of trends, which should be robust to noise and disturbance, and (ii) Interpretation of trends to perform the fault diagnosis.

A formal framework was built by Cheung and Stephanopoulos [23] for the representation of process trends. Janusz and Venkatasubramanian [24] identified a comprehensive set of primitives by which any trend can be represented.

2.5.3.2 Quantitative Methods

Methods based on quantitative feature extraction are discussed in this section. These approaches solve the problem as a pattern recognition problem. In a pattern recognition problem there are usually a set of predetermined classes. Data points should be classified and assigned to one of those predetermined classes.

Statistical Techniques

In real process operation, systems are subject to random disturbances. In contrast to deterministic systems, future state of stochastic systems cannot be determined by past and present states and future control actions. Since these systems are under random influences, it might be logical to express these systems in a probabilistic setting. When the system is not faulty, the probability distributions are the ones that are assigned to normal operating conditions. Anytime an anomaly happens in the system, these distributions change. This change might be considered in the parameters of probability distribution. Hence the problem of fault diagnosis in these methods could be equivalent to detecting the changes in the parameters of a static or dynamic stochastic system. The design of on-line change detection algorithms was presented by Basseville and Nikiforov [25].

One of the first attempts to apply statistical approaches in on-line monitoring and change detection was represented by quality control problems (Shewhart [26]). Multivariate statistical techniques are used to compress data and reduce the dimensionality to essential data for easier analysis.

Principal component analysis (PCA) was first proposed by Pearson [27] and has been included in textbooks (Anderson [28] and Jackson [29]) and research papers (Wold [30]; Wold, Esbensen and Geladi [31]).

Originating from the work of Wold [32], partial least square (PLS) method, was further developed in 1980s. PLS methods similar to PCA methods are useful in reducing the dimensions of both process variables and quality control variables. PLS methods were applied by Macgregor *et al.* [33] for process monitoring and diagnosis.

These methods approach the fault diagnosis problems as quality control problems. They perform classification by using knowledge of a priori class distribution. For example in Bayes classifier, density functions of the respective classes are used.

Neural Networks

Neural networks have become more popular in fault diagnosis recently. Because of their learning and interpolation capability they have allowed for different implementations over various processes. The applications of neural networks in fault diagnosis can be classified into two dimensions: (i) the architecture of the neural network such as back propagation, radial basis function, etc., and (ii) the learning algorithm such as supervised and unsupervised.

Back-propagation networks are the most popular supervised learning neural networks. Ungar, Powell, and Kamens [34] and Hoskins, Kaliyus, and Himmelblau [35] were

among the researchers who applied these networks for the problem of fault diagnosis. Fan, Nikolaou, and White [36] incorporate functional inputs in addition to normal inputs to improve the performance of the neural network for fault diagnosis. Incorporation of knowledge in neural network framework was discussed by Farrel and Roat [37] for better diagnosis. Integration of feed-forward neural networks with recurrent neural networks was proposed by Tsai and Chang [38]. Becraft and Lee [39] considered integration of neural networks with the expert systems for fault diagnosis.

2.6 Intelligent and Learning-based Methods for FDIR

Recently, intelligent methods have received increasing attention in the field of fault diagnosis. The artificial intelligent methods such as fuzzy systems, neural networks and expert systems have the potential to learn the plant model from the input-output data or learn fault knowledge from past experience. These methods are an important extension to the quantitative model-based approaches for residual generation. The learning-based approaches to fault diagnosis are methodologies for detecting, isolating and accommodating faults in nonlinear dynamical systems. The idea is to sense any malfunction in the system by using online approximation structures and adaptive nonlinear estimation methods. There are two main limitations in the use of learning methods for fault diagnosis. The first limitation is computational complexity and the second is derivation of analytical results on the performance properties of the fault diagnosis scheme. The first problem is overcome by advances in computer technology. However, derivation of analytical results on the fault diagnosis performance is difficult due to nonlinear nature of the systems that are involved.

A learning-based approach using neural networks for fault diagnosis was suggested in [40] in which the stability properties of fault diagnosis scheme were investigated for the restricted case of abrupt faults within modeling uncertainties. In [41], some robustness and sensitivity properties of the fault diagnosis were investigated. Sensitivity refers to the fault detection scheme ability to correctly determine the existence of a fault. Robustness with respect to model uncertainty may result in less sensitivity to faults. Hence in general, there is a tradeoff between robustness and sensitivity to fault. This tradeoff is usually avoided by assuming that the effect of faults is decoupled from the effect of modeling errors in the form of disturbances [42]. However, this assumption is not usually valid for real practical situations. Thus in some reports some priori known bounds are assumed for modeling errors and the effect of faults are considered more significant than that of modeling errors.

An input-output formulation for nonlinear fault diagnosis using learning-based approaches was introduced in [43].

2.6.1 Fault Diagnosis using Neural Networks

Neural networks can be used for the purpose of fault diagnosis to avoid difficulties involved with mathematical models. However these methods have their own inadequacies and shortcomings.

Neural networks model the system as a black box. They do not demonstrate rules that are governing the operation of the system and do not provide an understanding of the system.

They also do not predict the behavior of the system under uncertain circumstances. The required training time and complexity of training algorithm are further limitations to

these approaches. Neural networks in which neurons are used as membership functions, such as radial basis functions (RBF) and B-spline do not generalize satisfactorily when they are presented with data outside the training space. On the other hand multilayer perceptron (MLP) tends to do a better generalization. These networks should be updated through online training. The neural network is not valid when an unknown fault condition appears, because it is not trained to classify this type of fault. In systems requiring online training, adaptive training algorithms should be used. Acquiring all faulty data for neural network training is not usually possible. Thus, unsupervised learning which uses a Kohonen network and the counter-propagation network [44] is necessary in order to classify the faults not known a priori.

2.6.2 Fault Diagnosis using Fuzzy Logic

Fuzzy systems are useful in situations in which measurements are not precise or their interpretation depends strongly on the context or human opinion. Supporting the direct integration of the human operator into the fault detection and supervision process is the main advantage of using fuzzy approach. Fuzzy decision makings are very similar to expert systems and supervisory control. The issue that remains a challenge is to express the nonlinear model by a number of linear models. The number of these linear models should be minimized. Since the identification method using fuzzy logic depends on a large number of variables, this optimization is not easy.

Recently, applications of fuzzy logic to fault diagnosis approaches have received increasing attention. Observers can be used to generate symptoms based on the estimation of the output from the system. The first methods applied fuzzy set theory to express cause-effect relations in expert systems. The introduction of fuzzy logic can

provide reliable fault diagnosis for industrial applications through improving the decision making. However, the difficulty arises in training of the algorithm in the inference mechanism where knowledge is hidden in large amounts of data and knowledge is embedded in trained neural network. A fuzzy feedforward neural network is used to extract rules from an existing data base. [45] uses fuzzy logic for residual generation which can be an important way of considering modeling uncertainty at decision making rather than residual generation.

2.7 AI-based Fault Diagnosis

Design of fault diagnosis schemes for spacecraft operations is a complicated procedure. In the last two decades, Artificial Intelligence (AI) techniques have been applied to the above problem to simplify design procedures and improve the reliability of space systems. AI systems can be classified to two groups: (i) rule-based (model-free) (ii) model-based. Rule-based systems are designed according to statistical intuition, and experience and knowledge of the experts encoded as a set of rules. However, gathering these rules according to which the system is working and ensuring the completeness of the set of these rules might be a difficulty in applying these methods. These techniques are suitable for decision making and planning as an advisory role. On the other hand, model-based diagnosis systems are more reliable and predictable. [46] outlines some approaches to residual generation using methods of integrating quantitative and qualitative system knowledge based upon AI techniques.

2.8 Discrete Event System (DES) Approaches to Fault Diagnosis

In most model-based approaches to fault diagnosis, a continuous variable model in the form of differential or difference equations is needed. However, in a lot of cases these detailed models are not needed. In these cases, discrete event system approaches provide enough information for fault diagnosis as well as convenience in the procedure. Discrete-event systems are expressed by two features: (i) a discrete state set and (ii) an event-driven state transition mechanism. Among DES-based approaches the fault-tree approach seems to be the most popular [47]. In these approaches an illustrative display of the system is provided which is easy to understand. However they have some shortcomings. In these approaches the timing and order of the fault cannot be determined. Furthermore, it is difficult to analyze the systems that go through more than one phase of the operation. In [48] the timed event sequence was generated by the DES under supervision is compared with a set of specifications for normal operation, called templates. Finite-state automata have also been used for fault diagnosis. Using finite-state automata simplifies the design of diagnostic systems. It is also straightforward to capture the ordering of events using finite-state automata.

State-based approach is another method used for fault diagnosis [49]. In this approach it is assumed that the state set of the system can be partitioned according to the failure status of the model. The problem is to use sensor measurements and determine the block of the normal/faulty partition that the state belonged to at the time the last measurement was received. In [50], [51] and [52] passive online diagnosis was studies using an event-based framework. Failure diagnosability is addressed in [51]. In this reference a failure is called diagnosable if it can be detected and isolated after a finite number of events

following the time it has occurred. In [53] the detectability of failures and design of "minimum alphabet observers" are studied. There are other discrete-event based approaches to fault diagnosis, such as using causal networks and relational models that are not mentioned here.

2.9 Methodology Chosen in This Thesis

In this thesis a model-based approach is chosen for the problem of fault diagnosis. An observer-based method is applied to solve this problem. In this observer-based method, residual signals are generated to provide information about the status of the system as being faulty or not.

In the next chapter the spacecraft attitude control system to be monitored by the FDI system, will be introduced.

Chapter 3

3 Spacecraft Attitude Control System

In this thesis, the fault diagnosis problem is defined in the spacecraft attitude control system. The establishment of the Attitude Control field goes back to the late 1950s. On October 4th, 1957, the first satellite, Sputink I, was launched into earth orbit. This field has remained a topic of interest for a number of researchers ever since.

Attitude control of a spacecraft is control of the angular motion to impart a specific angular position with respect to celestial bodies, the lines of force of magnetic or gravitational fields, called standard or baseline reference system [54]. The control system, which brings the spacecraft to the desired angular position, is called the attitude control system.

One of the main factors in choosing a control system in a spacecraft is the required accuracy for the attitude control system. A system, whose accuracy is several degrees, is obviously less complicated and less expensive than the one with several hundreds of a fraction of a degree.

In certain cases, the control system should be able to provide attitude control in more than three directions, when different systems with different degrees of accuracy are present on the spacecraft, or when more degrees of freedom are required on the

spacecraft. For example when the solar panels need to be oriented toward the sun, but the spacecraft has to be angularly positioned towards a certain point on the planet earth.

There are two ways to control the spacecraft attitude: i) spin stabilization where the gyroscopic stiffness against external disturbance torques is provided through rotation of all or parts of the spacecraft, and ii) three axis control through which all rotational degrees of freedom are controlled by using one actuator for each axis. In the second case there is always some coupling present between different axes due to dynamic effect or control torque actuator misalignment. This thesis is concerned with a three-axis control scenario.

Spacecraft attitude control using angular momentum wheels as actuators has been studied extensively [55]. Flywheels have also many applications in the field of energy storage systems for the production of electrical power. They provide energy as an excess to the energy provided by solar panels in the spacecraft and can be used when the satellite is in an eclipse period. The alternative role that these wheels can play makes them even more popular to be used in spacecraft control field.

In this chapter, the structure of the spacecraft attitude control system is described. The reference systems to be used are briefly explained. This follows by describing the spacecraft dynamics and related equations of motion in the chosen reference system. Based on the obtained spacecraft dynamics, a control system will be designed to stabilize the system as well as satisfy the desired performance specifications.

3.1 Attitude Reference Systems

A reference system must be used in order to investigate the spacecraft dynamics. However any reference frame could be valid to achieve this purpose, choosing an

appropriate reference frame can simplify the whole process. Finding a convenient reference frame simplifies the calculations and reduces algebraic transformations. Four reference frames are defined here [56].

3.1.1 Spacecraft Fixed/Body Reference Frame

The origin of this frame is located at the spacecraft center of mass. The directions of the axes are determined according to the spacecraft geometry. The x-axis and y-axis are along the solar panels and the z-axis is along the height. This reference frame is shown in Figure 3-1.

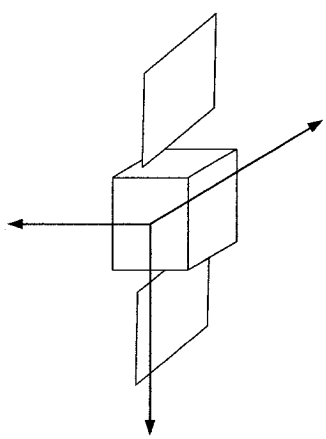


Figure 3-1: Spacecraft Fixed Reference Frame

3.1.2 Spacecraft Principal Axes Reference Frame

This reference frame is a body-fixed reference frame. Its origin is located at the spacecraft center of mass. The directions of the axes are along principal directions of the satellite body. The principal directions are the eigenvectors of the spacecraft inertia matrix. The dynamic equations can be expressed more conveniently in this reference frame [56].

3.1.3 Orbital Reference Frame

In this earth pointing reference system, the x-axis (Roll) is along the satellites velocity vector, the z-axis (Yaw) points towards the center of the earth and the y-axis (Pitch) completes the right-hand triad. As could be seen in the Figure 3-2, the angular position of spacecraft does not change the direction of the axes in this reference frame. The only factor that determines the direction of the axes is the position of the spacecraft's center of mass.

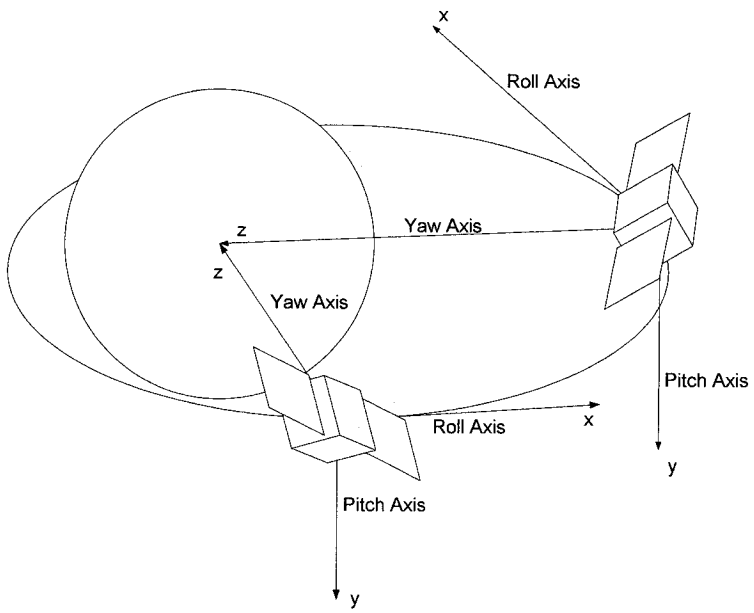


Figure 3-2: Orbital Reference Frame

3.1.4 Inertial Reference Frame

This earth-fixed inertial frame has its origin located at the earth's center. As is shown in Figure 3-3, in this reference frame, the x-axis points to the common point of Greenwich Meridian and the Equator, the z-axis in the same direction as the Earth's rotation axis, and the y-axis completes the right-hand triad.

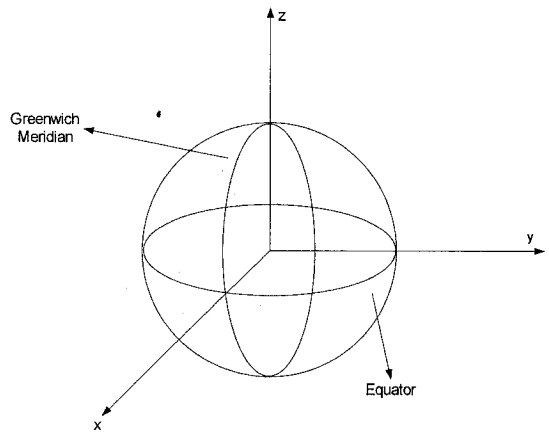


Figure 3-3: Inertial Reference Frame

The next step in the analysis of spacecraft attitude control system is the derivation of mathematical formula governing equations of the spacecraft body motion.

3.2 Spacecraft Dynamics

According to Newton's second law in a non-rotating coordinate system, the relationship between the torque $\vec{\tau}$ and the angular momentum $\vec{H}$ is:

$$\vec{\tau} = \dot{\vec{H}}$$

If the body reference is rotating with angular velocity ω as observed from the inertial reference frame, then this equation becomes:

$$\vec{\tau} = \dot{\vec{H}} + \omega \times \vec{H}$$

In this equation the cross product is due to change in magnitude of the components of H.

By definition of cross product:

$$\omega \times \vec{H} = (\omega_y H_z - \omega_z H_y)\vec{i} + (\omega_z H_x - \omega_x H_z)\vec{j} + (\omega_x H_y - \omega_y H_x)\vec{k}$$

This yields:

$$\vec{\tau} = (\dot{H}_x + \omega_y H_z - \omega_z H_y)\vec{i} + (\dot{H}_y + \omega_z H_x - \omega_x H_z)\vec{j} + (\dot{H}_z + \omega_x H_y - \omega_y H_x)\vec{k}$$

This could be rewritten in Euler's moment equations form as below:

$$\tau_x = \dot{H}_x + \omega_y H_z - \omega_z H_y$$
$$\tau_y = \dot{H}_y + \omega_z H_x - \omega_x H_z$$
$$\tau_z = \dot{H}_z + \omega_x H_y - \omega_y H_x$$

The angular momentum components and the angular velocity components could be related in the form below:

$$H_x = I_{xx}\omega_x - I_{xy}\omega_y - I_{xz}\omega_z$$
$$H_y = -I_{xy}\omega_x + I_{yy}\omega_y - I_{yz}\omega_z$$
$$H_z = -I_{xz}\omega_x - I_{yz}\omega_y + I_{zz}\omega_z$$

If the spacecraft body frame is considered aligned with the principal axes, the products of inertia become zero and this yields:

$$\begin{aligned} \tau_x &= \dot{\omega}_x I_{xx} + \omega_y \omega_z (I_{zz} - I_{yy}) \\ \tau_y &= \dot{\omega}_y I_{yy} + \omega_z \omega_x (I_{xx} - I_{zz}) \\ \tau_z &= \dot{\omega}_z I_{zz} + \omega_x \omega_y (I_{yy} - I_{xx}) \end{aligned} \qquad (3.1)$$

where x, y ,z now represent the principal axes of inertia. As can be seen in the above set of equations, there is coupling between different axes. The effect of this coupling becomes more significant when the angular speed values are large. After deriving the mathematical expressions for the equations of motion of the spacecraft dynamics the closed loop three-axis attitude control system can now be introduced.

3.3 Three-axis Attitude Control System

The main objective in this thesis is fault diagnosis and supervision rather than control. Hence, a classical controller which is able to stabilize the system and provide the desired output is satisfactory. In this section, first the closed loop attitude control system is described, which is shown in Figure 3-4. The model of the spacecraft is obtained

according to equations (3.1). The objective is to design a classical PD controller for this system. The actuator used to control the attitude of the spacecraft is a reaction wheel. In the upcoming section, reaction wheels are introduced in detail.

In Figure 3-4, the controller outputs pass through a saturation block. This is due to the fact that the reaction torque does not accept torque commands outside the interval [-5v, 5v]. The disturbances that affect the system before the model of the satellite are the combination of all external disturbances (namely, gravity-gradient, solar radiation, magnetic field, etc.). According to [57] all external disturbances have a magnitude of less than 10^{-4}. In this thesis, an aggregate representation of these disturbances is considered by a random signal with the magnitude of 10^{-3}.

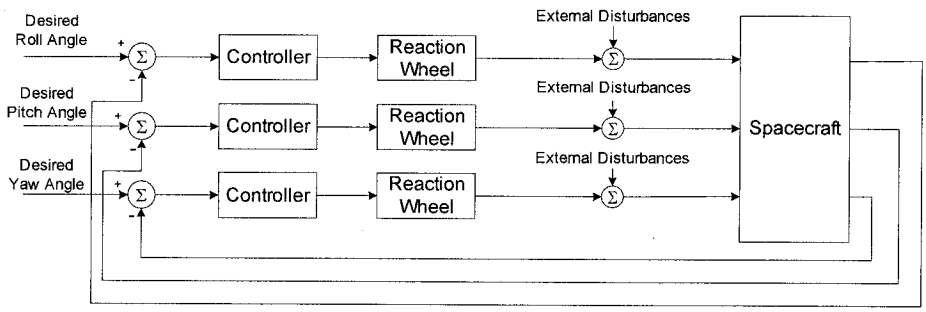


Figure 3-4: The three-axis stabilized spacecraft attitude control system

To design the controller, it is first assumed that there is no coupling between different axes. This is equivalent to assuming that the spacecraft is symmetrical around all the axes. This assumption eliminates the coupled terms in the set of equations (3.1) in page 32. Another assumption that is made in designing the controller is that the reaction wheel model is assumed to be a linear model which will be introduced later. When the performance of the PD control law in the single axis system with the simple reaction wheel model is guaranteed, then its validity for three-axis case and the nonlinear reaction

wheel model will be investigated. Figure 3-5 demonstrates the single axis control of the spacecraft.

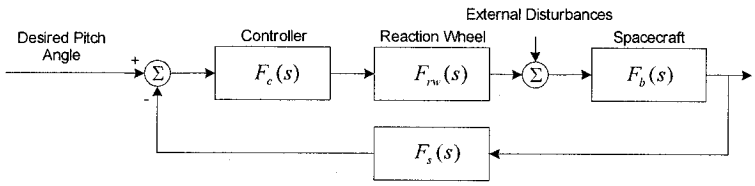


Figure 3-5: Spacecraft Single-axis Control

The transfer function from the input (desired angle) to the output (actual angle) will be:

$$T(s) = \frac{F_c(s)F_{rw}(s)F_b(s)}{1 + F_c(s)F_{rw}(s)F_b(s)F_s(s)}$$

where the transfer functions F_c, F_{rw}, F_b and F_s are defined according to Figure 3-5. In this system, $F_c(s)$ is the part that should be designed. Next, the models that will be used as the actuator, spacecraft dynamics and sensor will be specified explicitly.

3.4 Actuators

As was indicated earlier, the reaction wheels are assumed to be devices which provide the actuation signals for the spacecraft. In the next section the reaction wheels are introduced as the actuators employed in the ACS.

3.4.1 The Reaction Wheel

In this section, the model used for the reaction wheel is described. This model is a very accurate and detailed mathematical model that takes into account the effects of most disturbances which contribute to the spacecraft pointing jitter.

Reaction wheels are momentum exchange devices which provide reaction torque to a spacecraft and store angular momentum. A reaction wheel consists of a rotating flywheel, typically suspended on ball bearings by an internal brushless DC motor. A motor driver delivers commutated current to the motor proportional to an input voltage commanded by the user. Newton's laws can be used to model the reaction wheel mathematically. An almost ideal model for the reaction wheel is shown in Figure 3-6 [58]. In this fundamental block diagram the only loss that is considered is the friction torque.

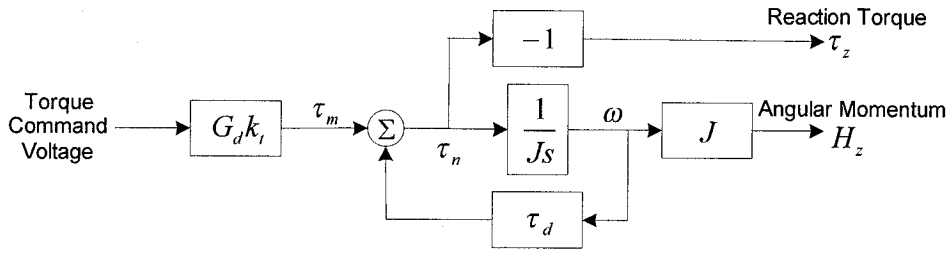


Figure 3-6: Nearly Ideal Reaction Wheel Model

The angular momentum stored in the flywheel, H_z is the product of the flywheel inertia J and the wheel speed ω, that is

$$\overline{H}_z = J\overline{\omega}$$

According to Newton's third law, the reaction torque applied to the spacecraft is equal and opposite to the net torque, τ_n, which accelerates or decelerates the flywheel. Also Newton's second law can be used to obtain τ_z, the reaction torque, that is

$$\tau_z = -\tau_n = -\frac{\partial H}{\partial t} = -\frac{\partial(J\overline{\omega})}{\partial t} = -J\frac{d\omega}{dt}$$

This yields:

$$\omega(t) = \frac{1}{J}\int \tau_n dt$$

According to Figure 3-6 the relationship between the net torque and the motor torque is:

$$\tau_n = \tau_m - \tau_d$$

The transfer function of this nearly ideal model will then be given by:

$$T(s) = \frac{JG_d k_t s}{Js + \tau_d}$$

Although studying the linear model of the reaction wheel is helpful in understanding and predicting its behavior, it cannot be a valid representation of the reaction wheel, applied in practice. Next, a more detailed model of the reaction wheel is introduced.

3.4.2 A High-Fidelity Model of the Reaction Wheel

In the previous subsection, a nearly ideal model of the reaction wheel was presented. Now a more detailed block diagram will be introduced to correspond to a better and more realistic model of the reaction wheel. This is shown in Figure 3-7 which relates to both ITHACO's standard Type A and Type B reaction wheels [58]. In this model additional blocks are included to model the performance beyond the nominal speed range, and as a function of temperature and bus voltage.

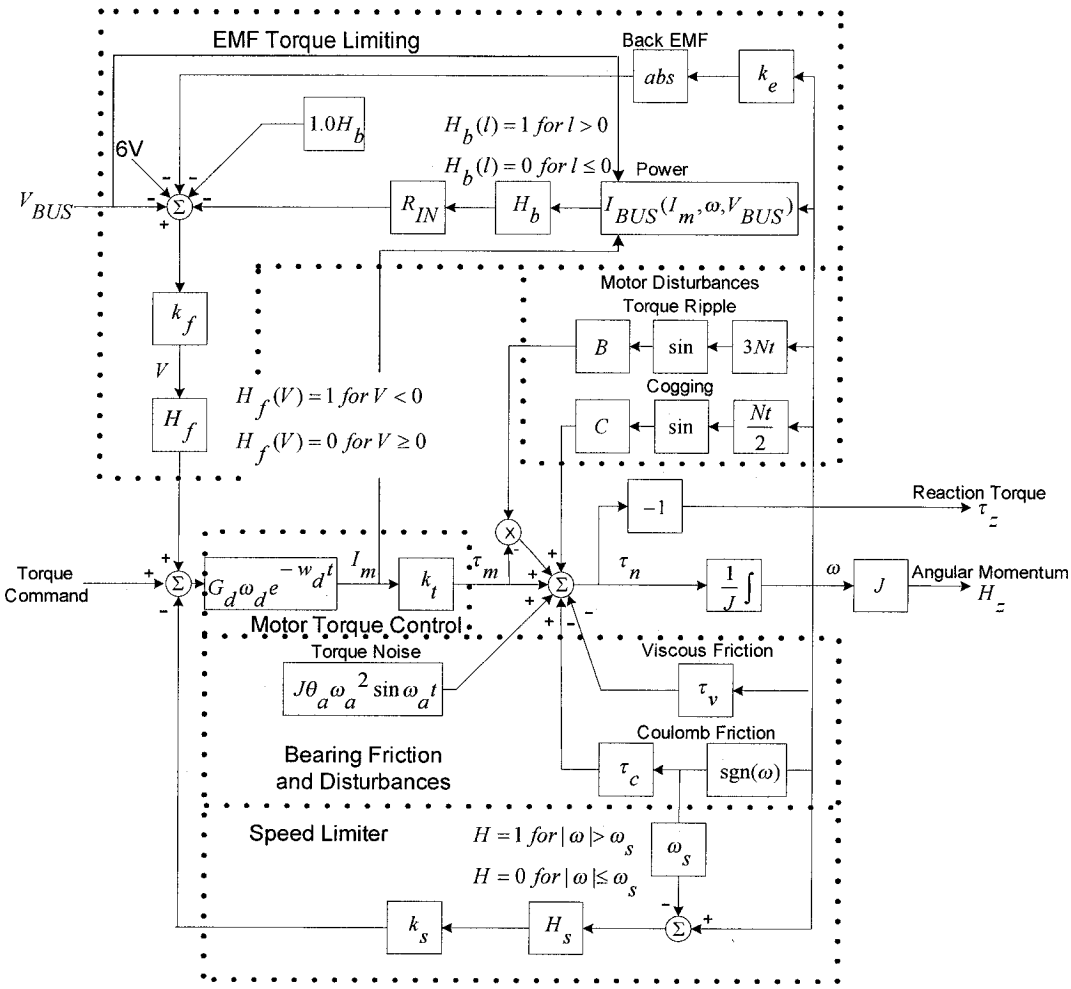


Figure 3-7: The High-fidelity Model of the Reaction Wheel [58]

The values of the parameters for different components and their units in this model can be found in Table 3.1. The constants in this table are considered for ITHACO's Type A reaction wheel [58].

Table 3-1: The Reaction Wheel Parameter Values [58]

Variable	Nomenclature	Units	Type A RWA
G_d	Driver Gain	A/V	0.19
ω_d	Driver Bandwidth (-3dB)	Rad/sec (Hz)	2000 (318)
k_t	Motor Torque Constant	N-m/A	0.029
k_e	Motor Back-EMF	V/rad/sec	0.029
k_s	Overspeed Circuit Gain	V/rad/sec	95
ω_s	Overspeed Circuit Threshold	Rad/sec (rpm)	690 (6600)
τ_c	Coulomb Friction	N-m	0.002
J	Flywheel Inertia	N-m-s^2	0.0077
N	Number of Motor Poles	-	36
B	Motor Torque Ripple Coefficient	N-m	0.22
C	Cogging Torque Amplitude	-	Zero
R_{IN}	Input Resistance	Ω	2.0
P_q	Quiescent Power	W	3.0
R_B	Bridge Resistance	Ω	2.0
	Torque Command Range	V	5
k_f	Voltage Feedback Gain	V/V	0.5
θ_a	Torque Noise Angle Deviation	rad (degrees)	0.05 (3)
ω_a	Torque Noise High Pass Filter Frequency	rad/sec	0.2

Motor Torque Control

The motor driver is essentially a voltage controlled current source with a gain G_d. It provides a motor current directly proportional to the torque command voltage. The motor has a torque constant, k_t, which delivers torque proportional to the input current, I_m.

Speed Limiter

A speed limiter circuit is applied to keep the speed of the flywheel in the safe range. This speed limiter measures the speed and provides it as a high gain negative feedback, k_s, into the torque command once a threshold for the speed, ω_s is exceeded. The Heavyside function H_s, enables the negative feedback when the wheel speed is larger than ω_s.

EMF Torque Limiting

When a low bus voltage occurs, the motor torque may be limited at high speeds due to increasing back-EMF, k_e, of the motor. I_{bus} is a nonlinear function of I_m, ω and V_{bus}. In case no power is being drawn from the bus, for instance during a deceleration when the energy is being removed from the flywheel, the Heavyside function H_b eliminates the voltage drop. In addition a reverse polarity protection diode drop of 1V dependent on H_b is included. The following formula [58] indicates how the bus current is affected by motor current, wheel speed and bus voltage.

$$I_{bus} = \left(\frac{1}{V_{bus} - 1}\right)(I_m{}^2 R_B + 0.04|I_m|V_{bus} + P_q + \omega I_m k_e)$$

Friction Model

The friction in the reaction torque is broken down into viscous friction, τ_v, which varies with temperature and speed, and coulomb friction, τ_c, which is a constant and its polarity varies depending on the wheel direction of rotation. The viscous friction is due to bearing lubricant. The lubricant viscosity is dependant on the temperature and this makes the viscous friction dependant on the temperature. The following formula [58] shows how the viscous friction is affected by the temperature:

$$\tau_v = (0.49 - \frac{0.02}{^\circ C}(T + 30^\circ C))\frac{mN - m}{rad/\sec}$$

The coulomb friction is caused by rolling within the bearings. This friction is independent of wheel speed. The polarity of this friction changes by the direction of wheel rotation.

Torque Noise

Torque noise is the very low frequency variation from the bearings, due to lubricant dynamics. This noise can have the most significant effect on the spacecraft pointing. It is largely a function of the lubricant behaviour. The torque noise can be approximated by the following formula [58]:

$$\tau_a = J\theta_a\omega_a^{\,2}\sin\omega_a t$$

Motor Disturbances

The torque motor in a reaction wheel can be a source of very high frequency disturbances due to the motor excitation and the magnetic construction. Brushless DC motors that are recently used in reaction wheel generate torque ripple at the commutation frequency, and cogging at a frequency corresponding to the number of poles and rate of rotation.

In the analytical model used in this thesis, all discontinuities have been replaced by one or a combination of sigmoidal functions in the form given below:

$$sigmoid(x) = \frac{1}{1+e^{\alpha x}}$$

In the model shown in Figure 3-7, the Heavyside functions H_b and H_f could be expressed as follows:

$$H_b(l) = \frac{1}{1+\exp(-\alpha l)}$$

$$H_f(l) = \frac{1}{1+\exp(\alpha l)}$$

The Sign and Absolute Value functions can be expressed as:

$$\text{sgn}(l) = \frac{1-\exp(-\alpha l)}{1+\exp(-\alpha l)}$$

$$abs(l) = \frac{1-\exp(-\alpha l)}{1+\exp(-\alpha l)} l$$

And finally the Heavyside function, H_s, can be written in the form shown below:

$$H_s(\omega) = \frac{1}{1+\exp(-\alpha(\omega-\omega_s))} + \frac{1}{1+\exp(\alpha(\omega+\omega_s))}$$

A formal mathematical representation of the reaction wheel model used in this fault diagnosis and isolation (FDI) problem may now be expressed as follows (for definition of states refer to Figure 3-7):

$$\begin{aligned} \dot{\omega} &= \frac{1}{J}[f_1(\omega) + k_t I_m[f_2(\omega)+1] - \tau_v\omega - \tau_c f_4(\omega) + n] \\ \dot{I}_m &= G_d\omega_d[f_3(\omega, I_m) - f_5(\omega)] - \omega_d I_m + G_d\omega_d r \end{aligned} \tag{3.2}$$

in which:

$$V(\omega, I_m, V_{bus}) = k_f \{V_{bus} - 6 - \frac{1}{1+\exp(-aI_{bus})}(1+R_{in}I_{bus}) - \frac{1-\exp(-ak_e\omega)}{1+\exp(-ak_e\omega)}k_e\omega\}$$

$$f_1(\omega) = C\sin\frac{Nt}{2}\omega$$

$$f_2(\omega) = B\sin 3Nt\omega$$

$$f_3(\omega, I_m, V_{bus}) = \frac{\exp(-aV(\omega, I_m, V_{bus}))}{1+\exp(-aV(\omega, I_m, V_{bus}))}V(\omega, I_m, V_{bus})$$

$$f_4(\omega) = \frac{1-\exp(-a\omega)}{1+\exp(-a\omega)}$$

$$f_5(\omega) = \frac{k_s[\omega - \omega_s f_4(\omega)]}{2}(\frac{1}{1+\exp(-a(\omega-\omega_s))} + \frac{1}{1+\exp(a(\omega+\omega_s))})$$

where f_1 and f_2 are functions due to the motor disturbances, f_3 is derived from the EMF torque limiting block, f_4 is a sigmoidal function replacing the sign function in the Coulomb friction, f_5 represents the speed limiter block, n is the torque noise, and r is the reference signal or torque command.

3.5 The Reaction Wheel Model used in Control Design

For the purpose of designing a control law, the nearly ideal model of the reaction wheel is applied. The transfer function for the wheel is given by:

$$F_{rw}(s) = \frac{JG_d k_t s}{Js + \tau_d}$$

Once the desired performance specifications are satisfied, the controller is tested for the closed loop system consisting of the high-fidelity and nonlinear model of the reaction wheel.

3.6 Body dynamics

Assuming a decoupled system for the spacecraft, the body dynamics transfer function for each axis will be:

$$F_b(s) = \frac{1}{I_{yy}s^2}$$

Here it is assumed that the controller is designed for the pitch axis. This could be done for all other axes as well.

3.7 Sensor

The sensor is supposed to measure the spacecraft angle with respect to the associated axis. For simplicity the transfer function considered for this block will be considered as a unity gain.

$$F_s(s) = 1$$

3.8 Controller

Designing the controller for the spacecraft could be a very complicated process. However, since the main purpose in this thesis is to design a fault diagnosis system, the satellite is assumed to be in normal operating conditions, namely Earth pointing. This assumption will eliminate complicated spacecraft maneuvers from the scope of this section and a classical PD controller can become an appropriate candidate. Hence the transfer function for the controller will be:

$$F_c(s) = k_p + k_d s$$

Therefore, the problem will be that of finding the proper controller gains which will be discussed next.

Control Design

Using the models discussed earlier, the closed loop transfer function is obtained as given bellow:

$$T(s) = \frac{(k_p + k_d s)\dfrac{JG_d k_t s}{Js + \tau_d}\dfrac{1}{I_{yy}s^2}}{1 + (k_p + k_d s)\dfrac{JG_d k_t s}{Js + \tau_d}\dfrac{1}{I_{yy}s^2}} = \frac{JG_d k_t k_d s + JG_d k_t k_p}{JI_{yy}s^2 + (\tau_d I_{yy} + JG_d k_t k_d)s + JG_d k_t k_p} \qquad (3.3)$$

The reaction wheel parameters may be found in Table 3.1. The moments of inertia are assumed to be:

$$I_{xx} = 18\frac{Nm}{rad/s^2}$$
$$I_{yy} = 16\frac{Nm}{rad/s^2}$$
$$I_{zz} = 21\frac{Nm}{rad/s^2}$$

Different values are assumed for the moments of inertia with respect to different axes in order to take into account the nonlinear effects of the spacecraft due to the coupling terms. However, the controller is first designed only for a single axis and then its performance is subsequently evaluated when generalized to the three-axis model.

The steady-state error for a step input with amplitude A and the transfer function in (3.3) is given by:

$$e_{ss}(\infty) = \lim_{s\to 0} sE(s) = \lim_{s\to 0}[s\frac{A}{s}(1 - T(s))] = 0$$

The desired percentage overshoot and settling time are considered to be 20% and 200 seconds, respectively. The characteristic polynomial in the transfer function (3.3) can be expressed in the standard second order form as:

$$s^2 + \frac{(\tau_d I_{yy} + JG_d k_t k_d)}{JI_{yy}}s + \frac{JG_d k_t k_p}{JI_{yy}} = s^2 + 2\zeta\omega_n + {\omega_n}^2$$

The parameters k_p and k_d are chosen in such a way that the above desired system performance specifications are satisfied. The PD controller that is chosen to meet these criteria is now obtained as:

$$F_c(s) = k_d s + k_p = 1000s + 100$$

Since the PD controller is guaranteed to solve the above design problem for a linear model, the performance result is only shown for the application of this control to the system including the nonlinear model of the reaction wheel.

Figure 3-8 shows the pitch angle for a single axis spacecraft system to demonstrate the achieved acceptable performance of the controller. Since the input reference signal, applied to each axis, represents the desired variation in the angle in that specific direction, it is limited to the interval of $[-\pi, \pi](rad)$. If the reference command is zero, it means that no change in the angle around the associated axis is required and if π or $-\pi$ is applied, it implies that the spacecraft should rotate 180 degrees around that axis in positive or negative direction, which is the maximum change in the angle variation.

In Figure 3-8 the desired variation in angle in the pitch direction is assumed to be $1\,rad$. As could be seen the settling time for the output of the system is less than $200\,\text{sec}$ and the overshoot is less than 20%, which means that the desired performance specifications are satisfied. The final value of the actual pitch angle is 1.002.

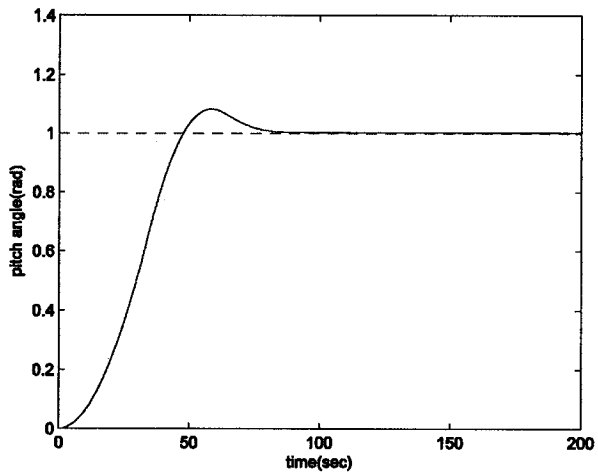


Figure 3-8: The simulated output of a single axis system with a linear control applied to a linear spacecraft model.

After validating the performance of the controller for a single-axis system, this controller is tested for the coupled 3-axis model. The results are shown in Figure 3-9. The maneuver that is considered for simulation in this case consists of rotations around different axes. The set of reference commands for this maneuver is assumed to be $\{2,\pi,1\}$ *rad* around roll, pitch and yaw axes respectively. Figure 3-9 demonstrates that all the desired angles could be achieved within the first 160 seconds. The final values of the actual roll, pitch and yaw angles are 2.005, 3.138 and 0.994, respectively.

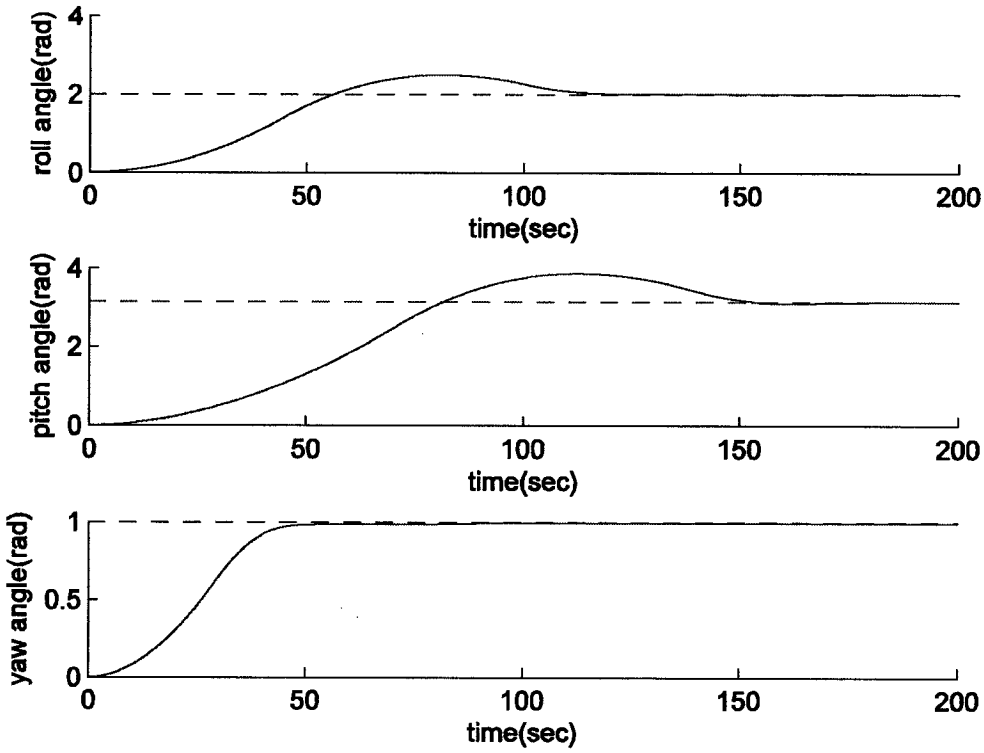


Figure 3-9: The simulated outputs of the three-axis system with linear control applied to nonlinear spacecraft model.

The next step in testing the designed control is to investigate its robustness against small changes in the system parameters. For instance the motor torque gain and the viscous friction gain are assumed to be subject to a 10% change at times 170sec and 180sec respectively. Figure 3-10 shows the simulated outputs for this scenario. As can be seen from this figure, the controller can compensate the effects of the disturbed parameters and force the outputs to converge to their desired values. As a result no deviation is seen in the outputs due to the uncertainty in the parameters. The final values of the actual roll, pitch and yaw angles are 2.006, 3.137 and 0.993, respectively.

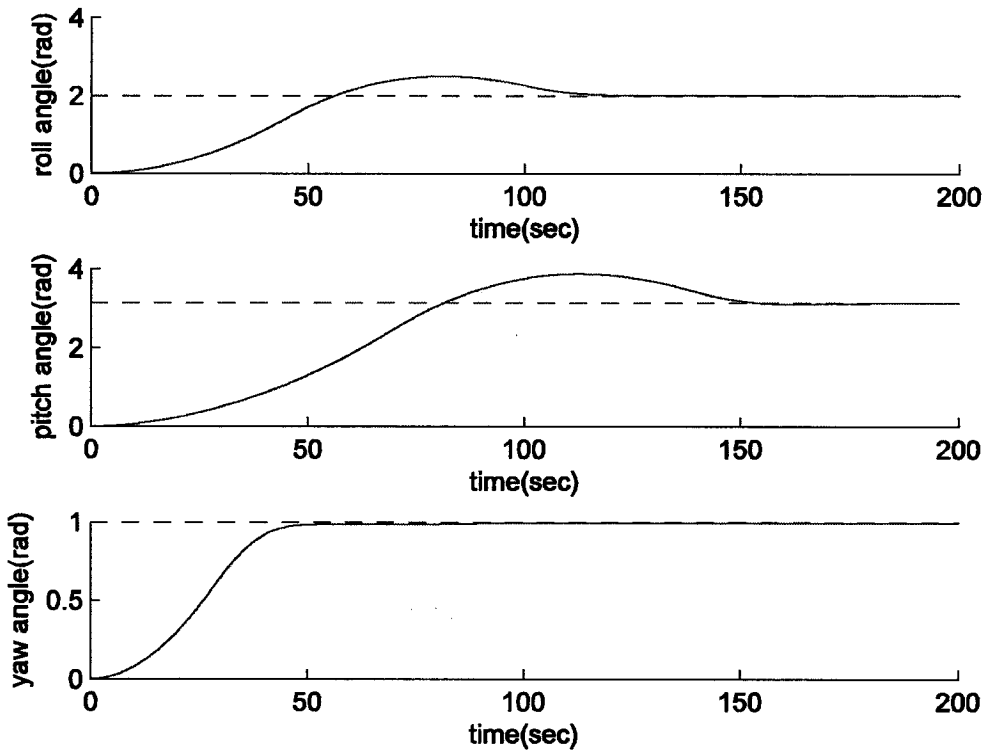


Figure 3-10: The simulated output of the three-axis control system in presence of parameter uncertainty

The final stage of our validation procedure is to verify the controller performance during different maneuvers. Figures 3-11 and 3-12 show the set of reference commands and the corresponding outputs when the angular position of the spacecraft changes several times. In Figure 3-11, the first desired spacecraft angular position is [2,2,1] (rad). This position is achieved in less than 150 seconds as shown in Figure 3-12. After 200 seconds, the spacecraft is required to change its angular position to [3,1,2] (rad). This time, the maneuver is performed in less than 100 seconds. Finally, the spacecraft is commanded to be situated in the new angular position of [1,0,3] (rad) which can be reached in less than 150 seconds as depicted in Figure 3-12.

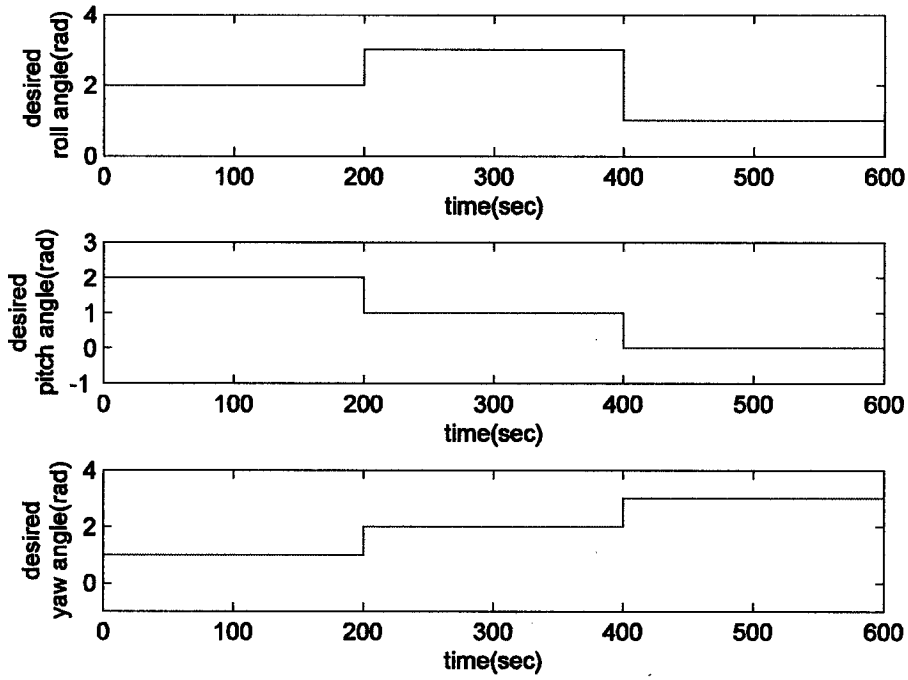


Figure 3-11: The set point reference commands to schedule different maneuvers for the spacecraft

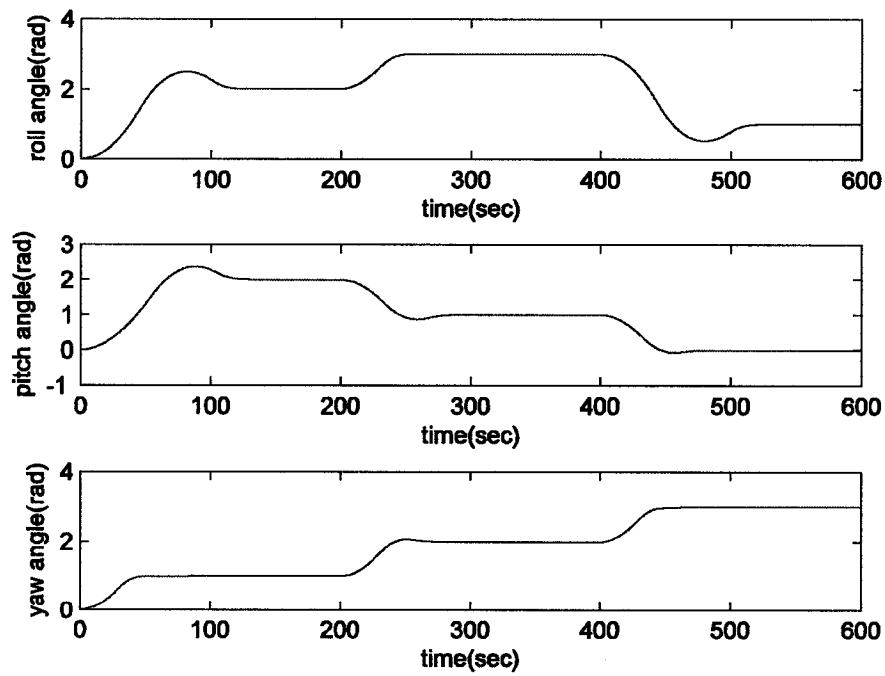


Figure 3-12: The actual outputs of the spacecraft following the desired reference commands

3.9 Conclusion

Now that the required spacecraft control is achieved, the problem of fault diagnosis can be investigated. Specifically, the spacecraft control system that is introduced in this chapter provides the framework in which the fault diagnosis algorithm will be tested and simulated in the following chapters.

Chapter 4

4 The FDI Architecture Design

The traditional approach to fault diagnosis is "hardware redundancy". In this approach multiple sensors, actuators and computers are applied and by polling from the hardware redundant components the decision about whether a fault in the system has occurred is made. However this approach is not economical, and in addition, more space and resources are required to accommodate the redundant components.

Considering the disadvantages of hardware redundancy, one may try to detect the fault in the system by data reconciliation or using the dissimilar measured values to cross check against each other. In other words, the diagnosis task could be achieved by using functional or analytical relationships between different measured variables in the system. This procedure is called "analytical redundancy". Consistency checking of measured variables in this scheme results in generating a signal called residual signal. This signal is supposed to be close enough to zero when there is no fault present in the system, and should distinguishably diverge from zero when a fault occurs in the system. A residual signal is usually obtained through a comparison between a signal in the system and its estimation in the mathematical model. Since analytical redundancy uses the mathematical model of the system to be monitored, it is referred to as model-based approach to fault

diagnosis. Thus model-based fault diagnosis can be defined as *the determination of faults of a system from the comparison of available system measurements with a priori information represented by the system's mathematical model, through generation of residual quantities and their analysis* [3].

The prominent advantage of the model-based approach to the hardware redundancy approach is the replacement of additional hardware with additional storage capacity and possibly greater computer power. Figure 4-1 depicts the concepts of analytical and hardware redundancy.

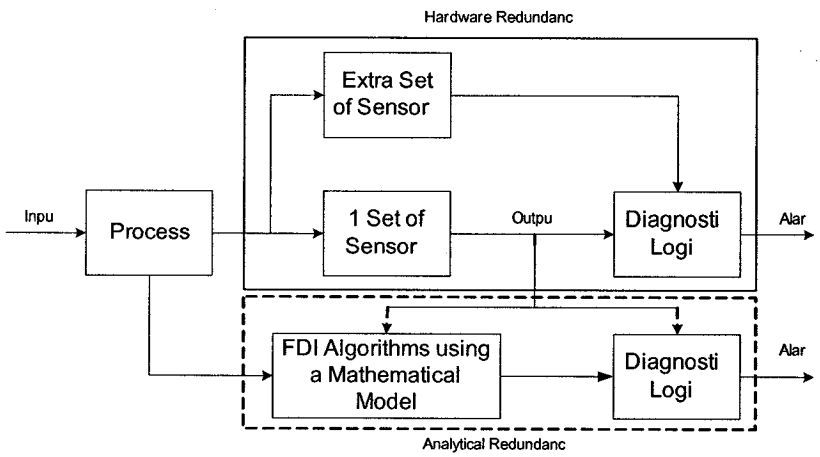


Figure 4-1: Hardware versus Analytical Redundancy

In model-based fault diagnosis, a threshold, either fixed or variable, is defined corresponding to the residual signal. When a sufficiently large fault occurs, the residual signal exceeds this threshold and the fault is detected. Usually there is more than one residual signal in the system and each of them is sensitive to a particular source of fault. Hence, the next concept of interest is to isolate the fault through monitoring the associated residual signals.

A model-based fault diagnosis system comprises of two stages of residual generation and decision making. Figure 4-2 shows this structure which was first suggested by Wilsky [16]. The description of these two stages is as follows:

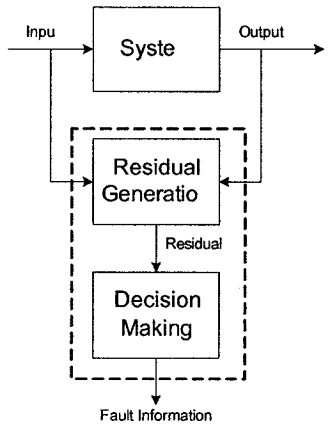


Figure 4-2: A two-stage procedure for fault diagnosis

1) Residual Generation: The purpose of this stage is to generate an auxiliary fault-indicating signal, called residual, which according to its value the faults could be detected. This signal should be normally close to zero when there is no fault and different than zero when a fault happens.

2) Decision-Making: The residuals are tested to determine the likelihood of existence of the fault in the system. This test could be based on a threshold test for the residuals or their moving averages.

Fault Diagnosis Structure

Figure 4-3 shows the relationship between the fault diagnosis and the control loop. Depending on the fault diagnosis objective, the supervised part of the system can be changed.

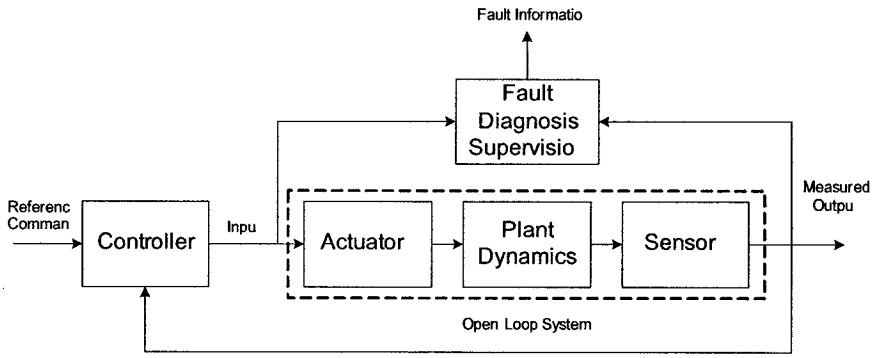


Figure 4-3: Fault Diagnosis and Control loop

In this thesis the faults considered are due and present in the actuator level. Hence, the supervision module is fed by the input and output of the actuator. The model-based approach that is chosen to address the problem of FDI is an observer-based method.

4.1 Observer-based Approaches

Observer-based techniques are quite popular model-based approaches to fault diagnosis. An observer is in fact an estimation of the model of the system to be monitored for the purpose of fault diagnosis. The outputs of this observer are estimations of the states of the system. The difference between the estimates of the system states and the actual states can be considered as an appropriate candidate for the residual vector.

In this section, the theory of state estimators will be briefly explained. A linear state estimator will be designed for the nearly ideal model of the reaction wheel. The performance of this observer will then be simulated as a fault diagnosis module.

4.1.1 State Estimators

In a lot of processes values of state variables are not available either because they are not directly reachable to be measured or there is a shortage of sensors to measure them. Since state vectors are needed for different purposes such as state feedback, it is fundamentally important to face this problem. What is going to be discussed in this section is how to estimate the unknown states by using the inputs, outputs and the relationship between them in a process. To achieve the objective of estimating the states of a system, a system is used which is called a state observer. State observers utilize the inputs and outputs of a system and as their outputs they generate an approximation of the system states. State observers could be classified into two groups in terms of the ratio of the states that they approximate. If in a process none of the states is available then a full-dimensional observer is needed. If only some states are unavailable a reduced-order observer is applied.

4.1.2 Full-dimensional State Estimator

Let us consider the following state and output representation of a system,

$$\begin{aligned} \dot{x} &= Ax + Bu \quad , x \in R^n, u \in R^m \\ y &= Cx \qquad\quad , \mathrm{y} \in \mathrm{R}^{\mathrm{q}} \end{aligned} \tag{4.1}$$

where matrices A, B and C are the system matrices of appropriate dimensions, n is the number of states, m is the number of inputs and q is the number of outputs. In this system it is assumed that none of the system states is available for measurement. However, the matrices A, B and C are assumed to be known matrices. In other words, the system model is available but the states are not.

For this problem the first approach that may come to mind is to duplicate the system and provide it with the same input that is injected to the original system. In this case if the initial states of the original system are forced on the duplicate model, then the states of the estimator will follow the same trajectory as the original model. Hence, the problem will be how to determine the initial states of the system. Since the duplicate model does not use the output of the original model, this type of estimator is called the open-loop estimator. Figure 4-4 shows this type of estimator.

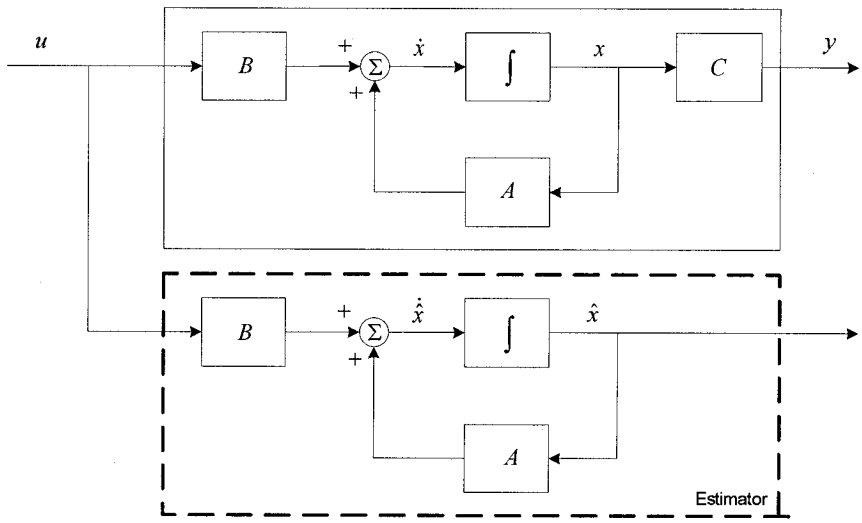


Figure 4-4: Open-loop state estimator

The open-loop estimator has some disadvantages. First, as was described earlier, the initial states should be computed every time the observer is switched on. Second, in real processes states and outputs are usually affected by noise and disturbances. This is not predicted in the open-loop observer and will cause at some occasions differences in the values of the system states and observer states. If the matrix A has an eigenvalue with a positive real part, this error will grow with time. All in all, it could be said that the open-loop observer is not a good choice.

Another way to build this observer is to convert the system state equations (4.1) to input-output differential equations. However, this approach is not an appropriate choice either. Because it is not easy to build differentiators of the input and output and even if that is not the problem, presence of differentiators makes the system very vulnerable to noise.

By taking a closer look at Figure 4-4 one may suggest to use the output of the original system as a source of information to improve the performance of the observer. This could be done by finding the error between the system output and the observer output and providing a feedback to the observer resulting in a closed-loop observer. This feedback has to go through the observer by a matrix gain L. The matrix L has to be found in a way such that the new observer becomes stable. Figure 4-5 illustrates the schematic of a closed-loop observer. It could be shown that if the state equations shown above are observable then the initial states can be found by the input and output of the original system. Thus if the system is observable its states could be estimated through this approach.

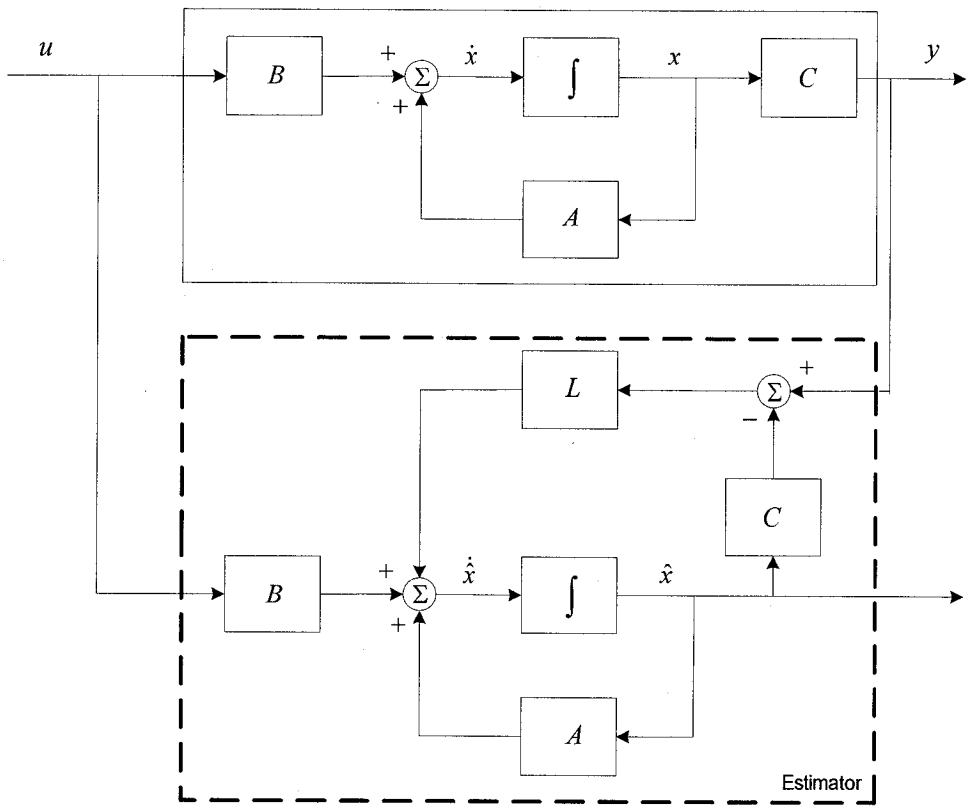


Figure 4-5: The closed-loop state estimator

The set of differential equations for the new observer according to Figure 4-5 is as follows:

$$\dot{\hat{x}} = A\hat{x} + Bu + L(y - C\hat{x}) \tag{4.2}$$

This vector equation could be rewritten as:

$$\dot{\hat{x}} = (A - LC)\hat{x} + Bu + Ly \tag{4.3}$$

Since the objective in designing an observer is to minimize the error between the states of the process and their estimates, an error vector is defined as:

$$e = x - \hat{x} \tag{4.4}$$

Using the above definition can result in expressing the vector equation (4.3) in the form below:

$$\dot{e} = (A - LC)e \tag{4.5}$$

If the location of the eigenvalues of the matrix A -LC can be controlled through proper choice of the matrix gain L, then the error vector can be controlled. If these eigenvalues are all forced to lie in the left-half plane, then the error vector converges to zero asymptotically. In this case even if there is an instantaneous discrepancy between the state and its estimate, due to disturbances or differences in initial state, the state estimate still follows the actual state by reducing that difference. Since in this closed-loop observer the error tends to zero asymptotically, this type of observer is called asymptotic state observer.

Since observers can estimate the states of a system, the difference between these estimations and the real values of the system states can be used to verify if the system is performing normally. However, there might be cases in which the system is under normal operating conditions but is affected by disturbances or model uncertainties. In these cases the observer should not classify the situation as an abnormal behavior of the system.

As described in the previous chapter, different observer-based approaches can be used for the purpose of fault diagnosis. In observer-based approaches the discrepancy between the system states and observer estimations can be a good criterion for detecting the presence of a malfunction in the system.

4.2 Fault Diagnosis using State Estimators

In this section a fault diagnosis algorithm is suggested for the spacecraft attitude control system. A state estimator plays the role of diagnosing the faults.

Figure 4-6 depicts the closed-loop attitude control system of a spacecraft together with the fault diagnosis scheme. The responsibility of diagnosing an anomaly or a fault is

played by three residual generators. Each residual generator monitors the performance of one of the reaction wheels and decides if that wheel is operating normally or if an anomaly or a fault has occurred in one of the components of that wheel. After detecting the occurrence of an anomaly or a fault, the source of the fault should be identified.

In this thesis, three common fault sources are considered for the reaction wheel. The first component, vulnerable to fault is the motor torque gain (k_t) which converts the current I_m to torque τ_c, the second is the viscous friction gain (τ_v), that is sensitive to temperature variations and the third source of fault is considered to be the bus voltage, V_{bus}.

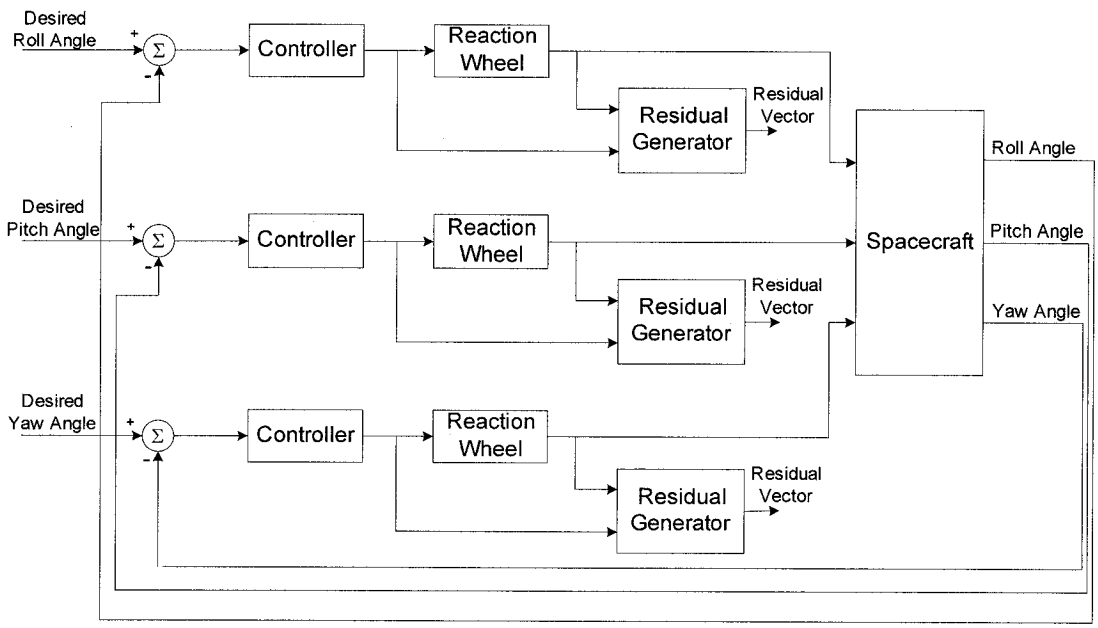


Figure 4-6: Spacecraft attitude control system and the FDI architecture

The first step in solving the FDI problem is to determine if a linear observer can be applied as the fault diagnosis module. To achieve this objective, the nonlinear model of the system is linearized. Then a linear state estimator is designed for the resulting

linearized system. The convergence properties of this observer are investigated. If the observer is convergent for some domain in which the system is operating, its performance will be evaluated as the fault diagnosis module. This performance becomes the benchmark for the nonlinear observer to be designed. Since a nonlinear observer has a higher degree of complexity, it is expected to execute the task of fault diagnosis better than the linear one. The advantages of using a nonlinear observer will be discussed in the following subsections.

4.2.1 Linear Observer Design

One way of designing a nonlinear observer is through designing a linear observer for the linearized model. Parameters of this linear observer can then be used to design a nonlinear observer.

In this thesis, a linear observer is first designed for the linearized version of the model introduced by the set of equations (3.2) in page 41. After designing the linear observer, the linear observer gain is used to design the nonlinear observer.

For the sake of simplicity the set of equations (3.2) in page 41 is rewritten in the form below:

$$\begin{aligned} \dot{\omega} &= \phi_1(\omega, I_m) + n \\ \dot{I}_m &= \phi_2(\omega, I_m) + G_d \omega_d r \\ y &= \begin{bmatrix} \omega \\ I_m \end{bmatrix} \end{aligned} \tag{4.6}$$

Based on the linearized state equations, the dynamics of a linear observer would be:

$$\begin{aligned} \dot{\hat{x}} &= A\hat{x} + L(y - \hat{x}) \\ \hat{y} &= \hat{x} \end{aligned}$$

where $\hat{x} = \begin{bmatrix} \hat{\omega} \\ \hat{I}_m \end{bmatrix}$. The matrix $A_{2\times 2}$ is the Jacobian matrix of the vector function $\begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix}$ with respect to the states ω, I_m at the operating point Q_0, namely

$$A = \left[\begin{bmatrix} \frac{\partial \phi_1}{\partial \omega} & \frac{\partial \phi_1}{\partial I_m} \\ \frac{\partial \phi_2}{\partial \omega} & \frac{\partial \phi_2}{\partial I_m} \end{bmatrix} \right|_{Q_0} \tag{4.7}$$

and where $L_{2\times 2}$ is the observer gain. The matrix L is chosen in such a way that the observer converges for the linear version of the model. In other words $A - L$ be a Hurwitz matrix.

An important point in this linearization is that the functions f_1 and f_2 should be ignored. These functions are modeling the motor disturbances and have the form shown below:

$$f_1 = C\sin(3N\omega t)$$
$$f_2 = B\sin(\frac{N\omega t}{2})$$

As can be seen the above two are time varying functions that will not allow for a simple way for linearization of the system. Hence, in the process of linearization the effects of motor disturbances in this design should be neglected. We have

$$\frac{\partial \varphi_1}{\partial \omega} = \frac{1}{J}[-\tau_v - 2\alpha\tau_c \frac{\exp(-\alpha\omega)}{(1+\exp(-aw))^2}]$$
$$\frac{\partial \varphi_1}{\partial I_m} = \frac{k_t}{J}$$
$$\frac{\partial \varphi_2}{\partial \omega} = G_d\omega_d(\frac{\partial f_3}{\partial \omega} - \frac{\partial f_5}{\partial \omega})$$
$$\frac{\partial \varphi_2}{\partial I_m} = G_d\omega_d\frac{\partial f_3}{\partial I_m} - \omega_d$$

where

$$\frac{\partial f_3}{\partial \omega} = \frac{\partial f_3}{\partial V}\frac{\partial V}{\partial \omega}$$

$$\frac{\partial f_4}{\partial \omega} = 2\alpha \exp(-\alpha\omega)/(1+\exp(-\alpha\omega))^2$$

$$\frac{\partial f_5}{\partial \omega} = \frac{k_s}{2}(1-\omega_s\frac{\partial f_4}{\partial \omega}(\frac{1}{1+\exp(-\alpha(\omega-\omega_s))}+\frac{1}{1+\exp(\alpha(\omega+\omega_s))})$$

$$\frac{k_s}{2}(\omega-\omega_s f_4(\omega))(\frac{\alpha\exp(-\alpha(\omega-\omega_s)}{(1+\exp(-\alpha(\omega-\omega_s))^2}-\frac{\alpha\exp(\alpha(\omega+\omega_s)}{(1+\exp(\alpha(\omega+\omega_s))^2})$$

$$\frac{\partial f_3}{\partial I_m} = \frac{\partial f_3}{\partial V}\frac{\partial V}{\partial I_{bus}}\frac{\partial I_{bus}}{\partial I_m}$$

$$\frac{\partial f_3}{\partial V} = \frac{\exp(-\alpha V)}{1+\exp(-\alpha V)} - \frac{\alpha\exp(-\alpha V)V}{(1+\exp(-\alpha V))^2}$$

$$\frac{\partial V}{\partial \omega} = k_f[-\frac{R_{IN}I_m k_e}{(V_{bus}-1)(1+\exp(-\alpha I_{bus})} - \frac{(1+R_{IN}I_{bus})I_m k_e \alpha\exp(-\alpha I_{bus})}{(1+\exp(-\alpha I_{bus}))^2} + k_e\frac{1-\exp(-\alpha k_e\omega)}{1+\exp(-\alpha k_e\omega)}$$

$$-\frac{\alpha {k_e}^2\omega\exp(-\alpha k_e\omega)}{(1+\alpha(-\alpha k_e\omega))^2}]$$

$$\frac{\partial V}{\partial I_{bus}} = -\frac{R_{IN}}{(1+\exp(-\alpha I_{bus})} - \frac{\alpha\exp(-\alpha I_{bus})(1+R_{IN}I_{bus})}{(1+\exp(-\alpha I_{bus}))^2}$$

$$\frac{\partial I_{bus}}{\partial I_m} = \frac{1}{V_{bus}-1}(2I_m R_B + 0.04V_{bus}\frac{1-\exp(-aI_m)}{1+\exp(-aI_m)} + k_e\omega$$

The error vector generated by the observer is defined as the difference between the actual states and their estimates, that is

$$e = \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} \omega - \hat{\omega} \\ I_m - \hat{I}_m \end{bmatrix} \tag{4.8}$$

The following differential equation gives the error vector dynamics:

$$\dot{e} = (A-L)e$$

By proper choice of L, the elements of the vector e can be made to exponentially converge to zero. Since these variables are sufficiently close to zero when there is no

fault present, they could be used as criteria for checking if there is a fault in the system. Hence, the above error vector will be used to generate the residual vector.

However, convergence of an observer is not a sufficient condition to justify its capability for fault diagnosis. In addition, in this thesis, the objective is to design a fault diagnosis algorithm for the nonlinear model as the linear observer estimates are less likely to converge to states of a nonlinear model. With this in mind, a nonlinear observer is designed to improve the possible inefficiencies of a linear fault diagnosis observer.

4.2.2 Nonlinear Observer Design

Since the linear observer does not have sufficient complexity to provide detailed information on the health or anomaly of the nonlinear reaction wheel, a nonlinear observer will also be designed.

The nonlinear observer could be obtained using the linear observer gain in the form shown below:

$$\begin{bmatrix} \dot{\hat{\omega}} \\ \dot{\hat{I}}_m \end{bmatrix} = \begin{bmatrix} \frac{1}{J}[f_1(\hat{\omega}) + k_t \hat{I}_m [f_2(\hat{\omega}) + 1] - \tau_v \hat{\omega} - \tau_c f_4(\hat{\omega})] \\ G_d \omega_d [f_3(\hat{\omega}, \hat{I}_m) - f_5(\hat{\omega})] - \omega_d \hat{I}_m \end{bmatrix} + L \begin{bmatrix} \omega - \hat{\omega} \\ I_m - \hat{I}_m \end{bmatrix} \qquad (4.9)$$
$$\hat{y} = \begin{bmatrix} \hat{\omega} \\ \hat{I}_m \end{bmatrix}$$

Using notations in (4.6), the set of equations (4.9) can be rewritten in a compact form shown below:

$$\begin{bmatrix} \dot{\hat{\omega}} \\ \dot{\hat{I}}_m \end{bmatrix} = \begin{bmatrix} \phi_1(\hat{\omega}, \hat{I}_m) \\ \phi_2(\hat{\omega}, \hat{I}_m) \end{bmatrix} + L \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} \qquad (4.10)$$
$$\hat{y} = \begin{bmatrix} \hat{\omega} \\ \hat{I}_m \end{bmatrix}$$

With appropriate definitions for ϕ_1 and ϕ_2, subtracting the state equations of the observer in (4.10) from the state equations of the system in (3.2) and using the notations in (4.6) and (4.8) yields:

$$\begin{bmatrix} \dot{e}_1 \\ \dot{e}_2 \end{bmatrix} = \begin{bmatrix} \phi_1(\omega, I_m) - \phi_1(-e_1 + \omega, -e_2 + I_m) \\ \phi_2(\omega, I_m) - \phi_2(-e_1 + \omega, -e_2 + I_m) \end{bmatrix} - L \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} \tag{4.11}$$

Next, the convergence of this nonlinear observer is investigated.

4.2.3 Convergence Properties of the Nonlinear Observer

In order to apply the nonlinear observer for the purpose of fault diagnosis, first convergence properties of this observer are investigated. Let's assume the matrix L is obtained through linearization of the system around the operating point:

$$Q_0 = \begin{bmatrix} \omega^* \\ I_m{}^* \end{bmatrix}$$

The Jacobian matrix of the vector function

$$\begin{bmatrix} \phi_1(-e_1 + \omega, -e_2 + I_m) \\ \phi_2(-e_1 + \omega, -e_2 + I_m) \end{bmatrix}$$

with respect to ω, e_1, I_m and e_2 is given by:

$$A_{nl} = \begin{bmatrix} \frac{\partial \phi_1}{\partial \omega} & \frac{\partial \phi_1}{\partial e_1} & \frac{\partial \phi_1}{\partial I_m} & \frac{\partial \phi_1}{\partial e_2} \\ \frac{\partial \phi_2}{\partial \omega} & \frac{\partial \phi_2}{\partial e_1} & \frac{\partial \phi_2}{\partial I_m} & \frac{\partial \phi_2}{\partial e_2} \end{bmatrix} \Bigg|_{\substack{\omega = \omega^* \\ e_1 = 0 \\ I_m = I_m{}^* \\ e_2 = 0}}$$

Now the set of equations (4.11) can be approximated as:

$$\begin{bmatrix}\dot{e}_1\\ \dot{e}_2\end{bmatrix}\approx\begin{bmatrix}\phi_1(\omega,I_m)\\ \phi_2(\omega,I_m)\end{bmatrix}+\begin{bmatrix}-\phi_1(\omega^*,I_m{}^*)\\ -\phi_2(\omega^*,I_m{}^*)\end{bmatrix}+\begin{bmatrix}-A_{nl12} & -A_{nl14}\\ -A_{nl22} & -A_{nl24}\end{bmatrix}\begin{bmatrix}e_1\\ e_2\end{bmatrix}+\begin{bmatrix}-A_{nl11} & -A_{nl13}\\ -A_{nl21} & -A_{nl23}\end{bmatrix}\begin{bmatrix}\omega-\omega^*\\ I_m-I_m{}^*\end{bmatrix}$$
$$-L\begin{bmatrix}e_1\\ e_2\end{bmatrix}$$

Considering the definition in (4.7), this yields:

$$\begin{bmatrix}\dot{e}_1\\ \dot{e}_2\end{bmatrix}\approx(A-L)\begin{bmatrix}e_1\\ e_2\end{bmatrix}+\begin{bmatrix}\phi_1(\omega,I_m)-\phi_1(\omega^*,I_m{}^*)\\ \phi_2(\omega,I_m)-\phi_2(\omega^*,I_m{}^*)\end{bmatrix}+\begin{bmatrix}-A_{nl11} & -A_{nl13}\\ -A_{nl21} & -A_{nl23}\end{bmatrix}\begin{bmatrix}\omega-\omega^*\\ I_m-I_m{}^*\end{bmatrix} \tag{4.12}$$

Let us write the equation (4.11) in the form shown below:

$$\begin{bmatrix}\dot{e}_1\\ \dot{e}_2\end{bmatrix}=\zeta(\omega,e_1,I_m,e_2)$$

in which $\zeta=\begin{bmatrix}\zeta_1\\ \zeta_2\end{bmatrix}$.

According to (4.12) in a close neighborhood of the operating point Q_0, the function ζ can be approximated as:

$$\zeta\approx(A-L)e+\theta(\omega,e_1,I_m,e_2)$$

where:

$$A-L=\frac{\partial\zeta}{\partial e}\Bigg|_{\substack{\omega=\omega^*\\ e_1=0\\ I_m=I_m{}^*\\ e_2=0}}$$

and $\theta=\begin{bmatrix}\phi_1(\omega,I_m)-\phi_1(\omega^*,I_m{}^*)\\ \phi_2(\omega,I_m)-\phi_2(\omega^*,I_m{}^*)\end{bmatrix}+\begin{bmatrix}-A_{nl11} & -A_{nl13}\\ -A_{nl21} & -A_{nl23}\end{bmatrix}\begin{bmatrix}\omega-\omega^*\\ I_m-I_m{}^*\end{bmatrix}$

Since the observer gain L is chosen such that $A-L$ be a Hurwitz matrix, then for any positive definite symmetric matrix Q, the solution P of the Lyapunov equation:

$$P(A-L)+(A-L)^T P=-Q$$

is positive definite. Now we introduce $V=e^T Pe>0$ as the Lyapunov function candidate. The derivative of this function along the trajectories of the error dynamics is:

$$\begin{aligned}\dot{V} &= e^T P\dot{e}+\dot{e}^T Pe = e^T P\zeta+\zeta^T Pe \\ &= e^T P[(A-L)e+\theta]+[(A-L)e+\theta]^T Pe \\ &= e^T[P(A-L)+(A-L)^T P]e+2e^T P\theta \\ &= -e^T Qe+2e^T P\theta\end{aligned}$$

Now it can be assumed that in some region Ω (where $(\omega,e_1,I_m,e_2)\in\Omega$) of the operating point we have the following condition:

$$\|\theta\|_2 \le \gamma\|e\|_2 \tag{4.13}$$

where γ is a constant. Therefore according to (4.13) for $(\omega,e_1,I_m,e_2)\in\Omega$:

$$\dot{V} < -e^T Qe+2\gamma\|P\|_2\|e\|_2^2$$

But $e^T Qe \ge \lambda_{\min}(Q)\|e\|_2^2$, where $\lambda_{\min}$ is the minimum eigenvalue of the matrix Q. Note that $\lambda_{\min}$ is real and positive since Q is symmetric and positive definite. Thus for $(\omega,e_1,I_m,e_2)\in\Omega$ we have:

$$\dot{V} < -[\lambda_{\min}(Q)-2\gamma\|P\|_2]\|e\|_2^2$$

Choosing $\gamma<(1/2)\lambda_{\min}(Q)/\|P\|_2$, ensures that $\dot{V}$ is negative definite. Therefore, the origin of the error vector is asymptotically stable. In other words, the observer states converge asymptotically to system states.

After constructing linear and nonlinear observers and investigating their convergence properties the next steps will be to define the residual signals and determine the fault decision making scheme. These steps will lead to addressing the fault diagnosis problem.

4.3 Fault Detection

The vector e generated by both linear and nonlinear observers has two components, namely: $e_1 = \omega - \hat{\omega}$ and $e_2 = I_m - \hat{I}_m$. In the case of convergence of observers these components converge to zero when there is no fault present in the system and the model of the system is close to the model of the observer. When the actual model is affected by faults, the observer estimates will have some discrepancy from the system states. Therefore, these non-zero error signals can represent good candidates for diagnosis of the faults. However, one may need to perform some refinements on these signals before using them as residual signals. These components may have very high frequency oscillations that will make the fault detection procedure quite difficult. To overcome this problem, the high frequency components of the error signals are removed by passing them through a low-pass filter. Therefore, the residual signal will be denoted as:

$$r = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}$$

where r_1 and r_2 are the low-pass filtered versions of e_1 and e_2, respectively.

After generating the filtered residual signals, the performance of the model is simulated for a large number of non-faulty scenarios to determine a threshold for each residual signal. The upper and the lower threshold bounds will be the maximum and minimum values that each residual signal can achieve respectively under normal operating conditions of the system. When the residual signals are within this interval, no

fault or anomaly is detected to be present in the system. Whenever at least one of the residual signal thresholds is exceeded, the FDI scheme issues a flag about the existence of a fault.

4.4 Fault Isolation

Basically for the purpose of fault isolation each residual signal in the residual vector should be decoupled from all the sources of fault but one. In our proposed approach, only two residual signals are available while there are three fault sources in the system. Therefore, it is not possible to uniquely identify all sources of faults correctly by simply observing the two residual signals. Because even in the best case that each type of fault triggers one residual signal, there is one residual signal which is sensitive to two faults. Thus further signal processing techniques should be applied to resolve this ambiguity. This problem is not pursued in this thesis.

4.5 Conclusions

In this chapter linear and nonlinear observers are designed to be applied as the fault diagnosis module. The convergence properties of these observers are formally investigated and the behavior of the nonlinear fault diagnosis scheme is described.

In the next chapter the performance of the nonlinear observer is investigated in the closed-loop ACS through simulation. The sensitivity of the residual signals to different types of faults will also be investigated.

Chapter 5

5 Comparative Analysis and Simulation Results

In this chapter, the performance of both linear and nonlinear observers as fault diagnosis modules are investigated and evaluated. Two cases will be considered. In the first case the degree of nonlinearity in the system is "low" and the linear observer is shown to be convergent. Therefore, this linear observer is tested to verify if it is a good candidate for solving the ACS fault diagnosis problem. The performance of this observer is investigated through simulations for different fault scenarios and is compared to that of a nonlinear observer. The advantages and motivations for selecting a nonlinear estimator are investigated. In the second case the degree of nonlinearity is "high" and the linear observer is shown to be non-convergent. Importantly enough, the nonlinear observer is shown to be convergent and be able to perform the fault diagnosis task. The performance and capabilities of applying this observer under this scenario will also be investigated.

We remind that in the following experiments, a random signal represents the external disturbances in Figure 3-4. The magnitude of this signal is assumed to be less than 10^{-3} *Nm*. According to table 3-1 the torque noise in Figure 3-7 is considered to be a sinusoidal signal in the form given below:

$\tau_a = 1.54 * 10^{-5} \sin 0.2t$

5.1 The Performance of a Linear Observer as a Diagnosis Module

The performance analysis of a linear observer is divided into two cases. Choosing the nominal bus voltage as 12v places the system in a "less" nonlinear area. This ensures that a linear observer is able to estimate the original nonlinear system. However, it will be shown that nevertheless it does not usually provide accurate information about the state of the system during the presence of faults. In other words, although it provides rather accurate estimates of the system states; it is not a good candidate for fault diagnosis.

5.1.1 First Scenario: High Bus Voltage, $V_{bus}=12V$

The first necessary condition that an observer should satisfy to be chosen as a fault diagnosis candidate is convergence. Figures 5-1 and 5-2 show the convergence of the error signals obtained by the linear observer corresponding to different initial states.

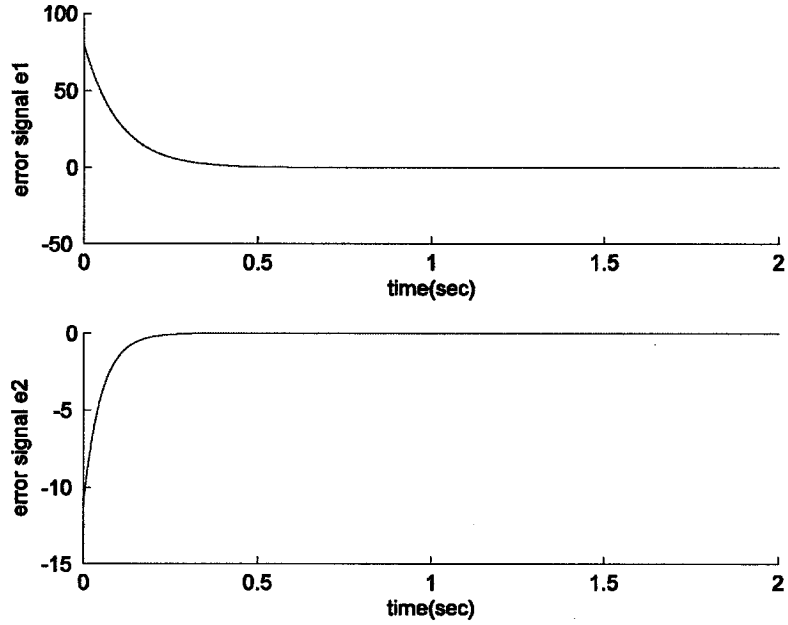


Figure 5-1: Convergence of the linear observer for high bus voltage with initial states $e_1(0)$=80 and $e_2(0)$=-12.

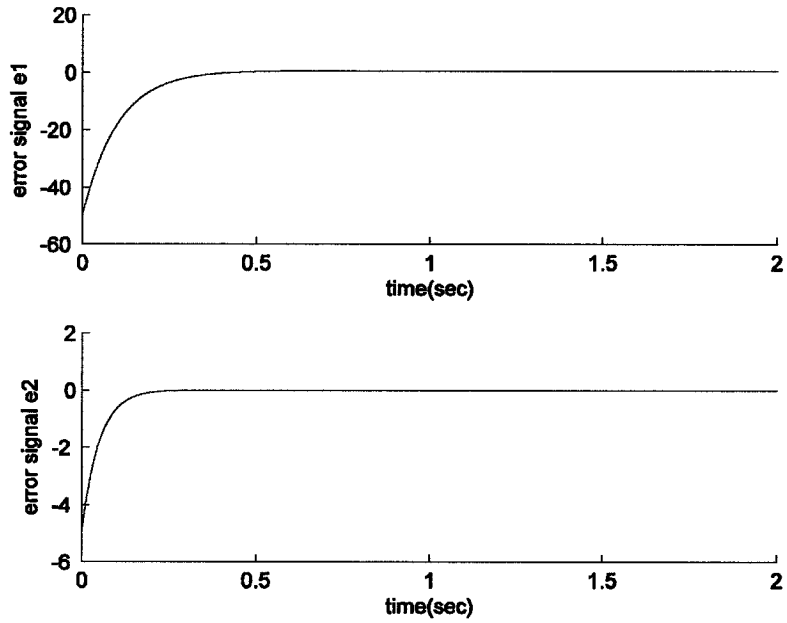


Figure 5-2: Convergence of the linear observer for high bus voltage with initial states $e_1(0)$=-50 and $e_2(0)$=-5.

Since this observer is clearly convergent for a high bus voltage, it could be tested for the fault diagnosis task. The first step in applying the linear observer as a fault diagnosis module is to define the residual signals.

As was pointed out in the previous chapter the components of the error vector generally contain some high frequency signals due to external noise and disturbances. Hence in order to use them as residual signals, as stated earlier, these signals are passed through a low-pass filter. The next step in the fault diagnosis algorithm is to define the "safe" (that is healthy) boundaries or intervals on the residual signals. This could be achieved by conducting different experiments in which the system is immune to any fault. These experiments can result in obtaining an upper and a lower boundary for the residual signals. After performing numerous simulations, the boundaries for the residual signals are obtained. In order to be somewhat conservative, the bounds are considered 5% farther from the maximum and minimum obtained from these experiments. Consequently, the residuals detect the presence of a fault whenever they exceed either their upper or lower boundary.

After determining the "safe" boundaries for the residuals, we investigate if these boundaries are violated when the system is faulty. Three types of faults will be considered. These faults occur either in the gain k_t, the gain τ_v or the bus voltage.

5.1.1.1 Faults in the Motor Torque Gain (k_t)

In this section, the performance of the linear observer is simulated for faults in the gain k_t. Figures 5-3, 5-4 and 5-5 show the residual signals r_1 and r_2 for three different sizes of faults in the gain k_t. All the faults are assumed to occur at time 400sec and

persist thereafter (permanent fault). The simulations are conducted for fault sizes of 10%, 40% and 70%. The dashed lines depict the "safe" or "healthy" boundaries that constrain the residual signal when there is no fault in the system. As could be seen in Figure 5-3 the residual signals are not sensitive enough to violate the safe band when a 10% fault occurs in the gain k_t. However, the residual signal r_1 shows some sensitivity to this fault. Figures 5-4 and 5-5 demonstrate the cases for 40% and 70% faults respectively. These faults are significant enough to push the residual signal r_1 out of its no-fault region. However, this residual signal does not stay outside this region for any of these cases. Returning of the residual signal r_1 to its no-fault zone implies that the fault is no longer present in the system which is a wrong conclusion of the situation presented by the linear observer. In all these cases the residual signal r_2 remains insensitive to these faults.

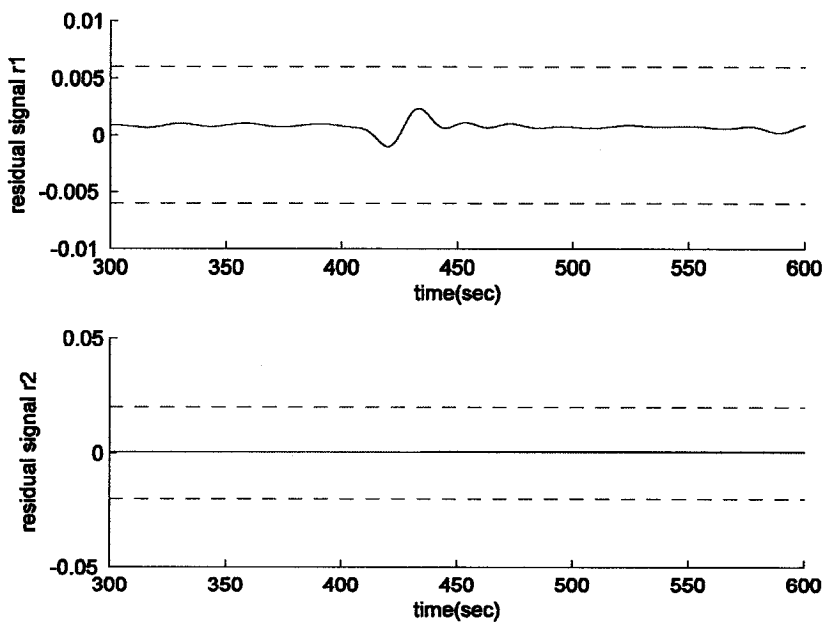


Figure 5-3: The residual signals due to a 10% fault in the gain k_t

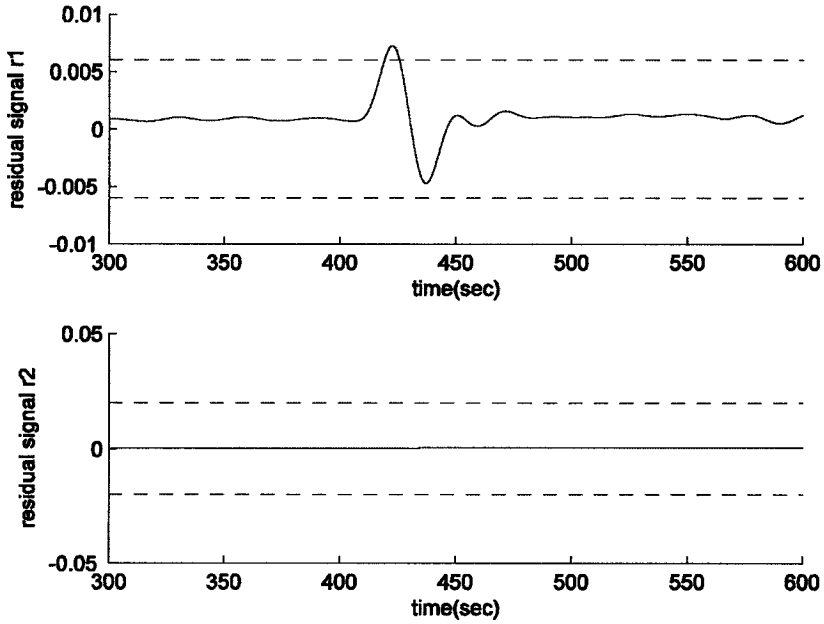


Figure 5-4: The residual signals due to a 40% fault in the gain k_t

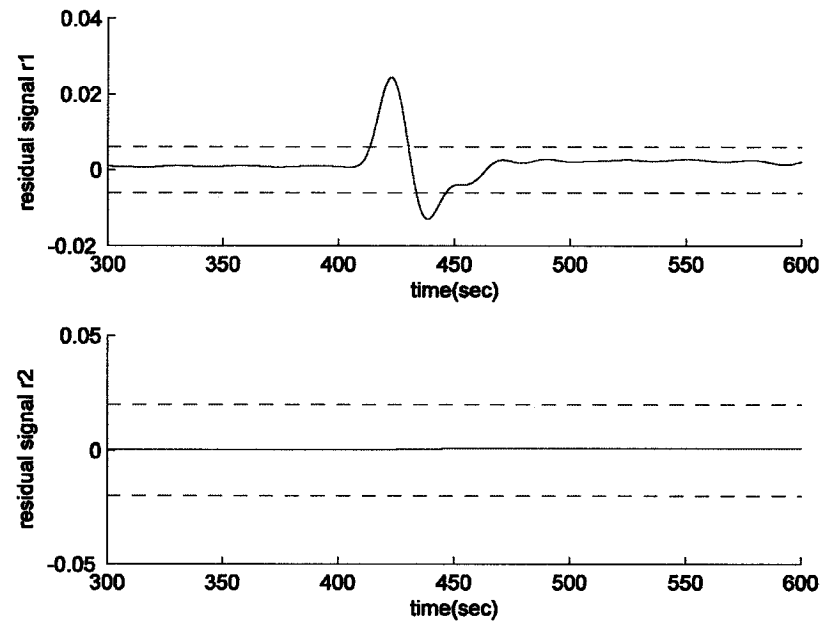


Figure 5-5: The residual signals due to a 70% fault in the gain k_t

5.1.1.2 Faults in the Viscous Friction Gain (τ_v)

The second source of fault that is considered in this analysis is the viscous friction gain. Figures 5-6, 5-7 and 5-8 depict the residual signals r_1 and r_2 for faults of different magnitudes in this gain. The time of occurrence of faults is assumed to be 400sec and the fault is assumed to persist thereafter (permanent anomaly). The fault sizes in these simulations are considered to be 20%, 50% and 80%. As could be seen in Figure 5-6, the residual signal r_1 is excited after the 20% fault happens. However, it does not cross any of the boundaries and settles down after about 100 seconds. Figures 5-7 and 5-8 depict the behavior of the residual signals for faults of 50% and 80% in the viscous friction gain. It can be seen that similar to the case of fault in the gain k_t, the residual signal r_1 breaks its boundaries but does not remain outside the no-fault zone to ensure persistence of the fault. In all these cases the residual signal r_2 is not disturbed by this fault. Therefore, the linear observer is not successful in detecting the faults in the viscous friction gain.

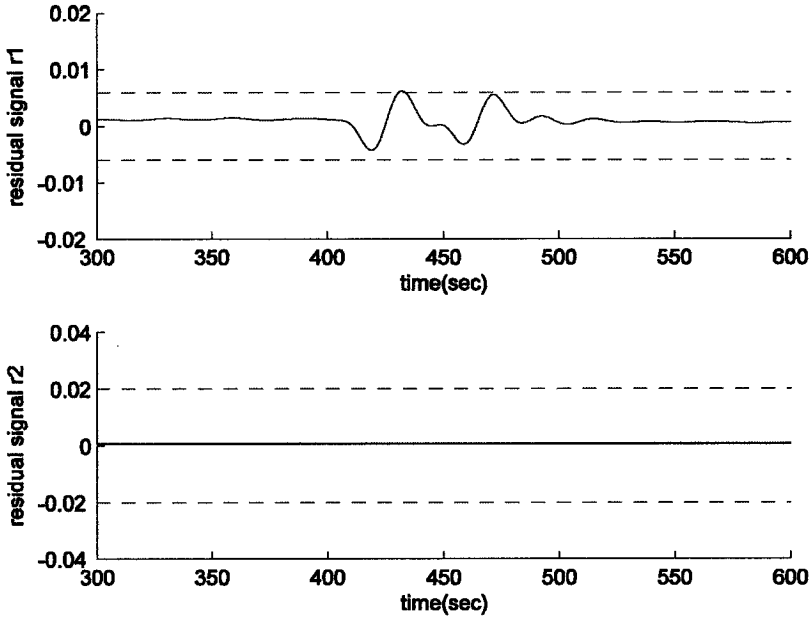


Figure 5-6: The residual signals due to a 20% fault in the gain τ_v

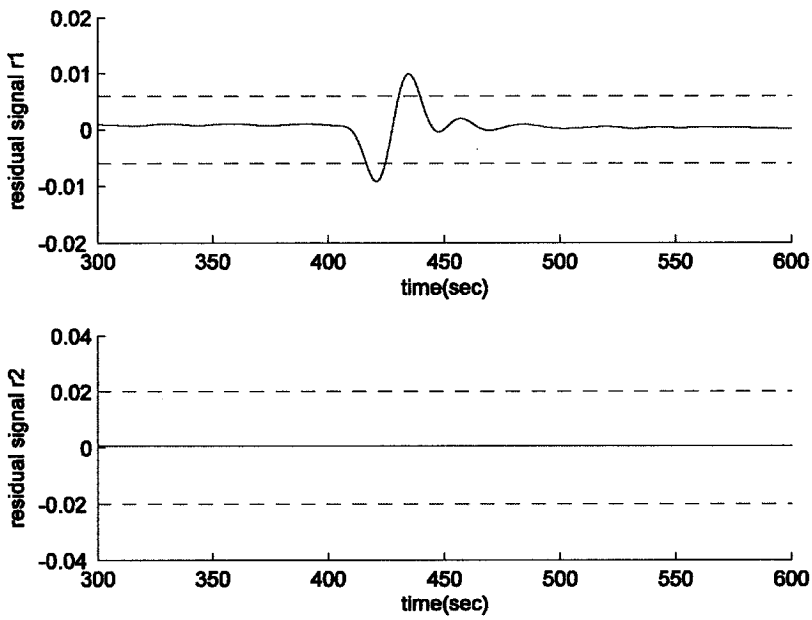


Figure 5-7: The residual signals due to a 50% fault in the gain τ_v

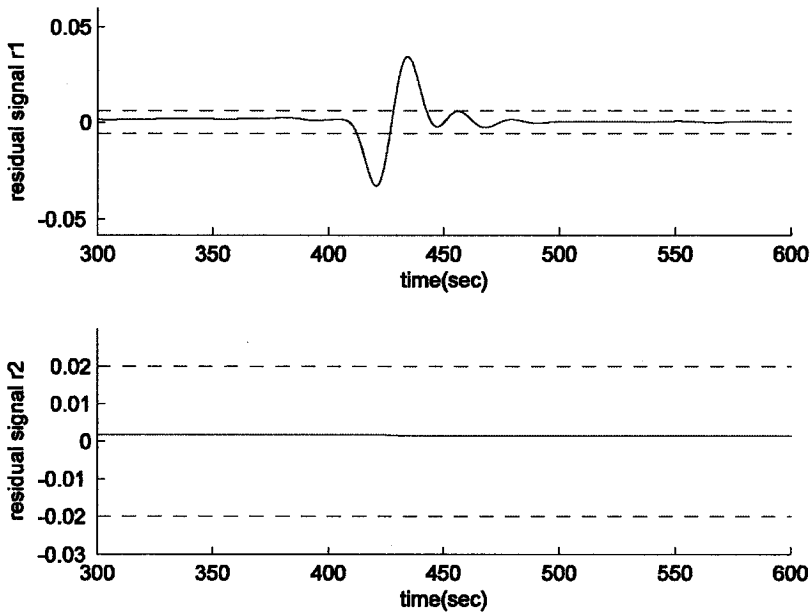


Figure 5-8: The residual signals due to an 80% fault in the gain τ_v

5.1.1.3 Fault in the Bus Voltage

The last source of fault discussed here is the bus voltage. In this section, the performance of the fault diagnosis algorithm confronting the linear observer is tested. Figures 5-9, 5-10 and 5-11 demonstrate the diagnostic signals when a fault happens in the bus voltage at 400sec and it remains thereafter (permanent fault). The fault sizes are 30%, 40% and 50% respectively. In this case both residuals escape from their safe region after a fault occurs in the bus voltage. It was seen in the two previous fault scenarios that the residual signal r_2 is not sensitive to faults in the gains k_t and τ_v. Hence, we can conclude that the excitation of the residual signal r_2 is equivalent to occurrence of a fault in the bus voltage. This implies that the linear observer can identify the faults in the bus voltage, no matter if it is the only fault affecting the system or if there are multiple faults present in

the system. The reason is that the residual signal r_2 is exclusively sensitive to faults in the bus voltage. In the next section the performance of the linear observer is investigated when an intermittent fault is present in the bus voltage.

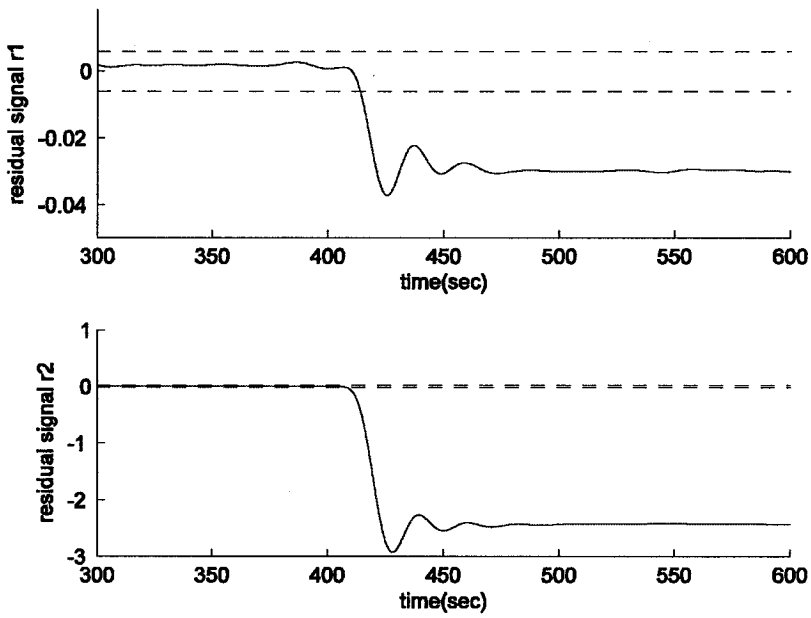


Figure 5-9: Detection of the 30% fault in bus voltage by the residual signals at time 414sec

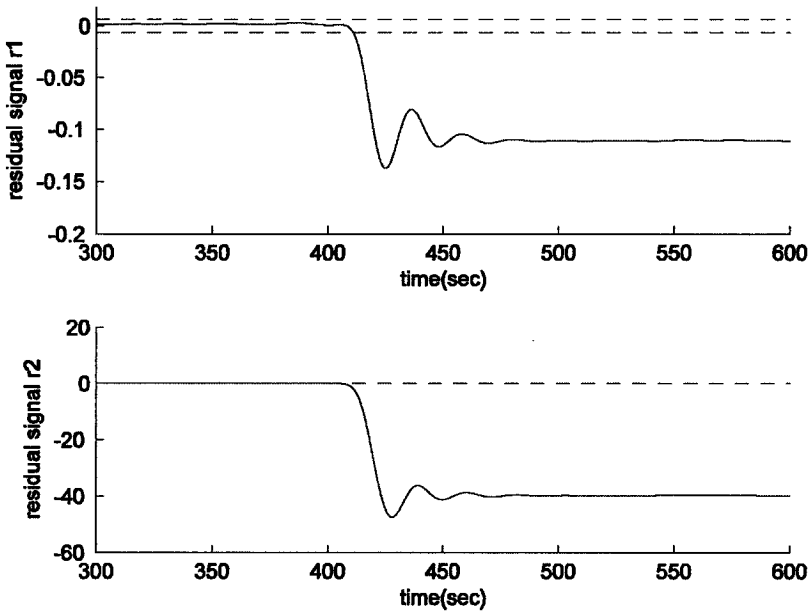


Figure 5-10: Detection of the 50% fault in bus voltage by the residual signals at time 411sec

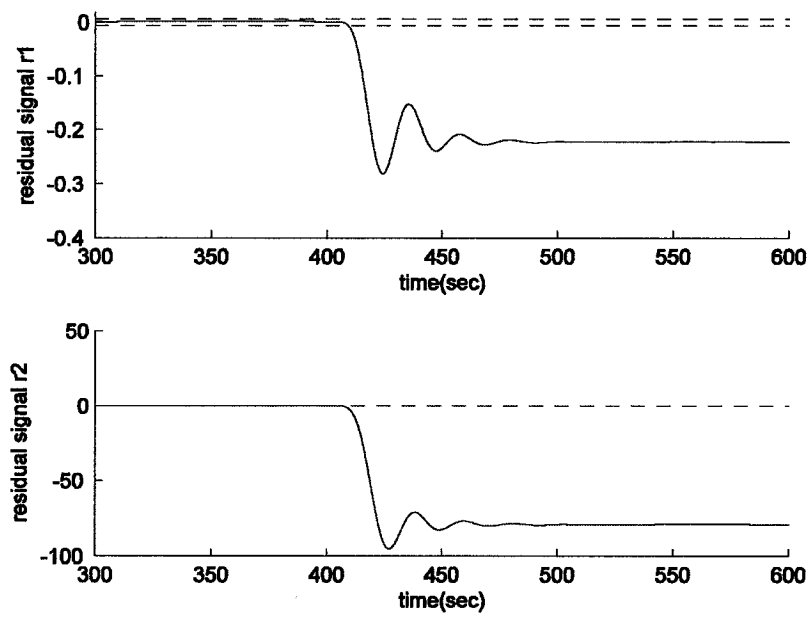


Figure 5-11: Detection of the 70% fault in bus voltage by the residual signals at time 409sec

5.1.1.4 Intermittent Fault in the Bus Voltage

Since the linear observer can detect the fault in the bus voltage, its performance should be validated when the faults in the bus voltage are not persisting. In other words, it should be investigated if the residual signals can still establish and determine when the fault vanishes. Since the linear observer cannot detect the faults in the gains τ_v and k_t, there is no point in investigating the case of intermittent faults in these gain.

Figure 5-12 shows the case when a 40% percent fault occurs in the bus voltage and this fault vanishes after 100 seconds. The residual signals both move out of their no-fault boundaries after the fault occurs. Interestingly enough, when the fault is removed these signals return to their no-fault zone. Therefore, the linear observer is still able to properly detect a fault when the fault in V_{bus} is temporary.

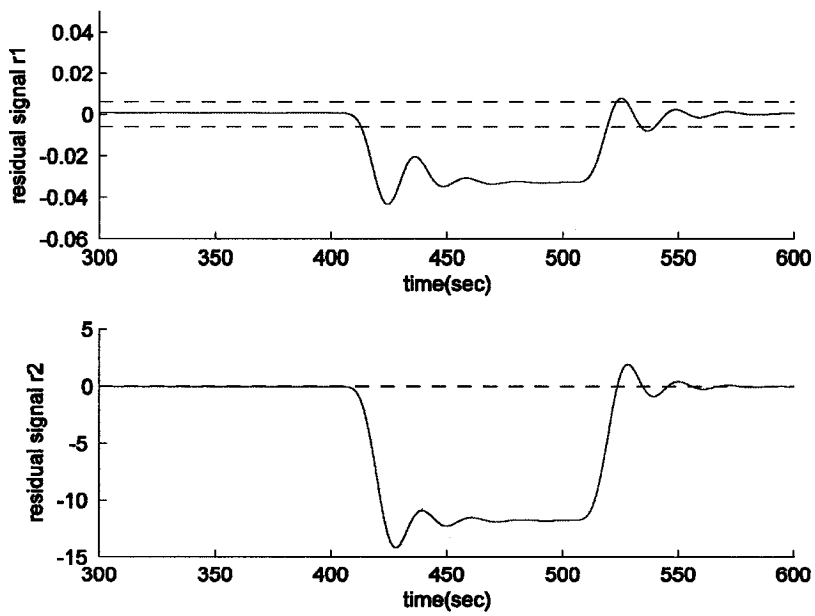


Figure 5-12: The residual signals due to an intermittent 40% fault in the bus voltage

In the next section the possibility of using the linear observer for low bus voltages is examined.

5.1.2 Second Scenario: Low Bus Voltage, V_{bus}=6V

Since the linear observer was convergent for a high bus voltage, its performance capability as a diagnosis module was investigated in the previous subsections. However, the estimates provided by a linear observer do not converge to the states of the original system when the system is working under the operating condition $V_{bus} = 6V$. Thus, the linear observer cannot be used as an appropriate choice for diagnosis in the low bus voltage scenarios. Figure 5-13 shows the error signals produced by this observer under this scenario. It can be seen that these error signals do not converge to zero and as a consequence the linear observer cannot be proposed to perform the fault diagnosis task.

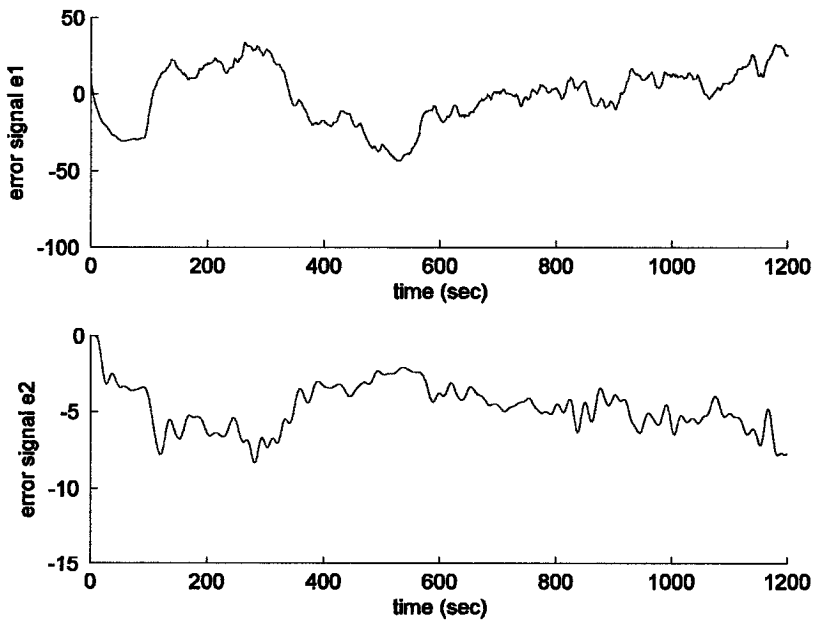


Figure 5-13: Divergence of the error signals under low bus voltage operating scenario

Next, the performance of a nonlinear observer, as designed in the previous chapter, will be investigated.

5.2 The Performance of a Nonlinear Observer as the Diagnosis Module

In this section a similar procedure as in the previous section is followed to analyze the performance of the nonlinear observer. This analysis comprises of two cases: (i) high bus voltage $V_{bus} = 12V$ and (ii) low bus voltage $V_{bus} = 6V$.

Figures 5-14, 5-15 and 5-16 show the error signals provided by the nonlinear observer. Since these signals converge to zero the observer is convergent and it may be used as a potential candidate for the purpose of fault diagnosis.

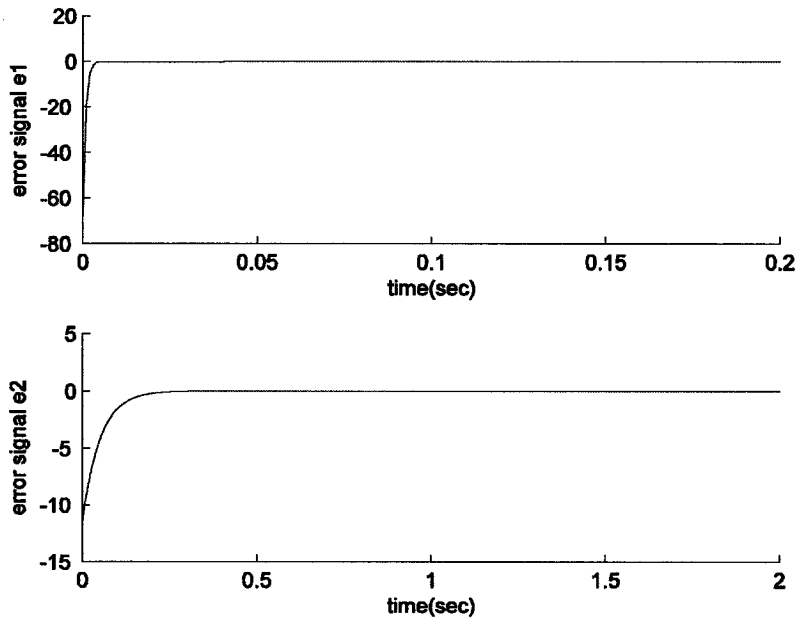


Figure 5-14: Convergence of the nonlinear observer for high bus voltage with initial states $e_1(0)$=-70 and $e_2(0)$=-12.

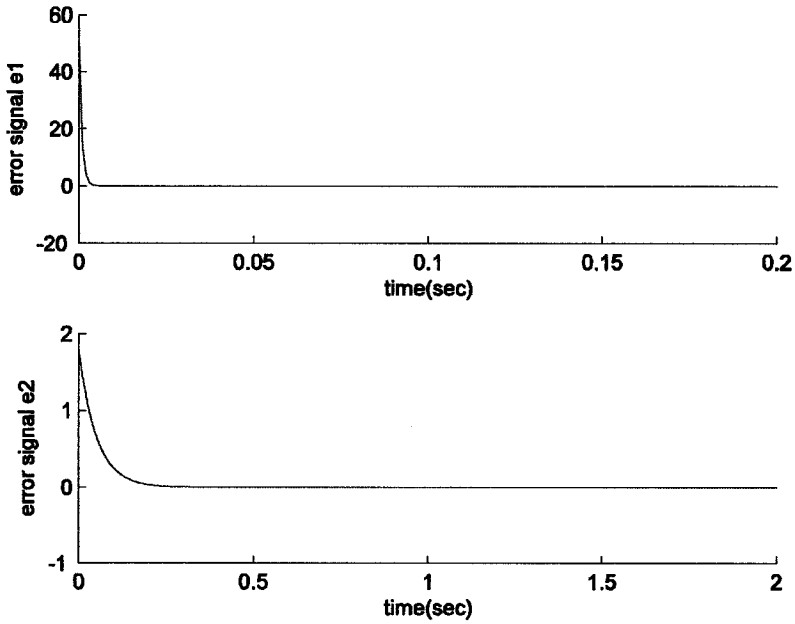


Figure 5-15: Convergence of the nonlinear observer for high bus voltage with initial states $e_1(0)$=50 and $e_2(0)$=2.

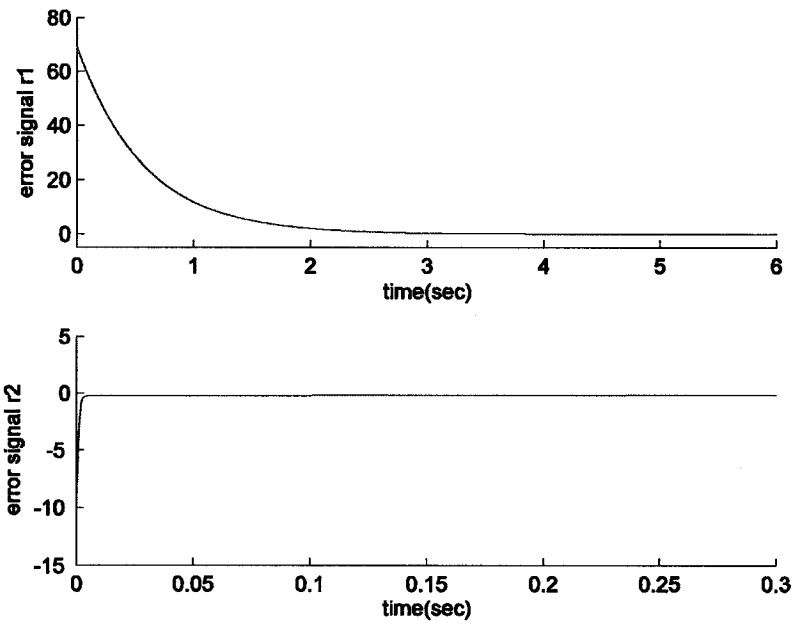


Figure 5-16: Convergence of the nonlinear observer for low bus voltage with initial states $e_1(0)$=70 and $e_2(0)$=-10.

The first step in applying the nonlinear observer as a fault diagnosis module is to define the residual signals. Similar to the linear observer the error signals generated by the nonlinear observer carry high frequency signals which can be eliminated by a low-pass filter. The next step is to define the "safe" or "healthy" boundaries on the residual signals. These boundaries are found by extensively simulating the system without having any faults affecting the system. These experiments lead one in obtaining a band for the residual signals r_1 and r_2. However, in the case of high bus voltage in all non-faulty scenarios, the residual signal r_2 remains very close to zero. Hence, in the case of a high bus voltage the signal r_1 detects a fault whenever it exceeds either its upper or lower boundary and the signal r_2, generates a warning message whenever it diverges from zero.

5.2.1 First Scenario: High Bus Voltage, V_{bus}=12V

It was shown earlier that the linear observer is convergent for high bus voltages. However, it could not detect the faults in the viscous friction and the motor torque gains. In this section, the capability of the nonlinear observer for compensating the shortcomings of the linear observer is investigated.

5.2.1.1 Faults in the Motor Torque Gain (k_t)

The first scenario considered here is to examine the performance of the nonlinear observer when a fault occurs in the gain k_t. Figures 5-17, 5-18 and 5-19 show the residual signals r_1 and r_2 for three different fault sizes in the gain k_t. All faults are assumed to happen at time 400sec and remain thereafter (permanent faults). The

simulations are conducted for fault sizes of 10%, 40% and 70%. The dashed lines depict the boundaries that constrain the residual signal when there is no fault in the system. As can be seen the residual signal r_1, exceeds the boundary, after the fault happens at 400sec. On the other hand, the residual signal r_2 shows no sensitivity to this type of fault. In these figures, it is obvious that when the fault becomes more significant, the residual signal r_1 goes farther away from its safe or no-fault region.

Another important point about the residual signal r_1 is that it can show the persistence of the fault as long as the fault is present. It was seen in Chapter 4, that the residual signal of the asymptotic observer for the nearly ideal model was sensitive to fault but it was not capable of ensuring the presence of the fault by a change in its steady state value. As a result of these experiments it could be stated that the faults in the gain k_t can be detected by the residual signal r_1.

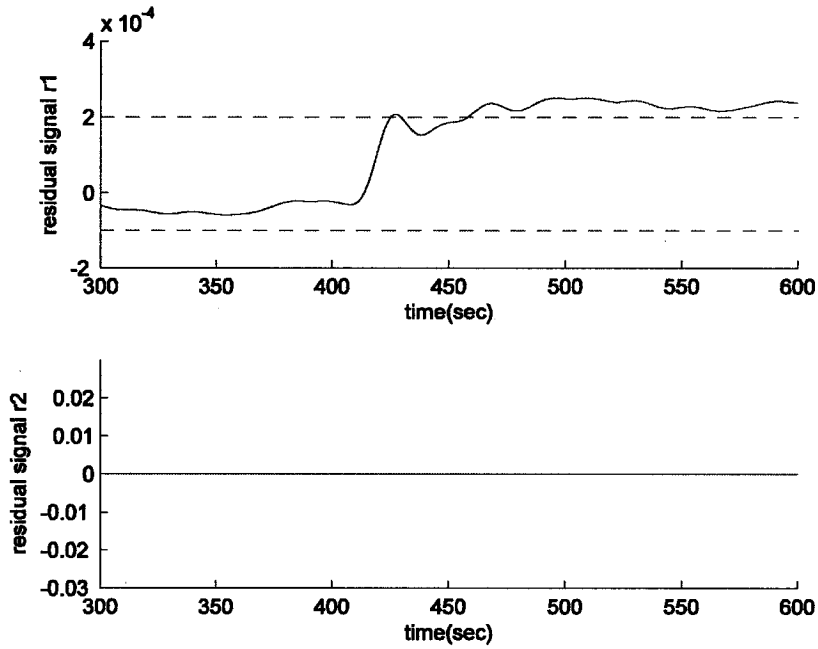


Figure 5-17: Detection of the fault at time 458sec by the residual signals generated due to a 10% fault in the gain k_t

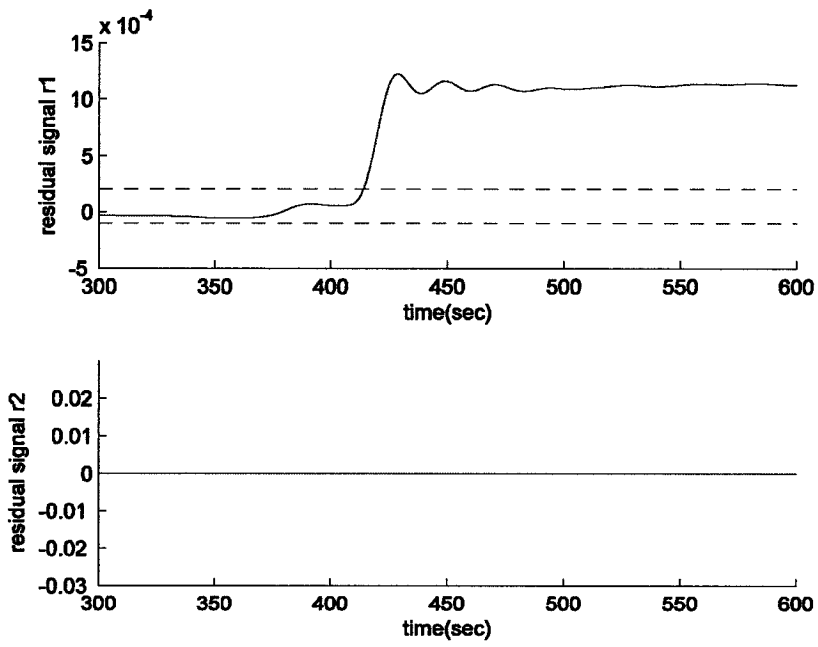


Figure 5-18: Detection of the fault at time 413sec by the residual signals generated due to a 40% fault in the gain k_t

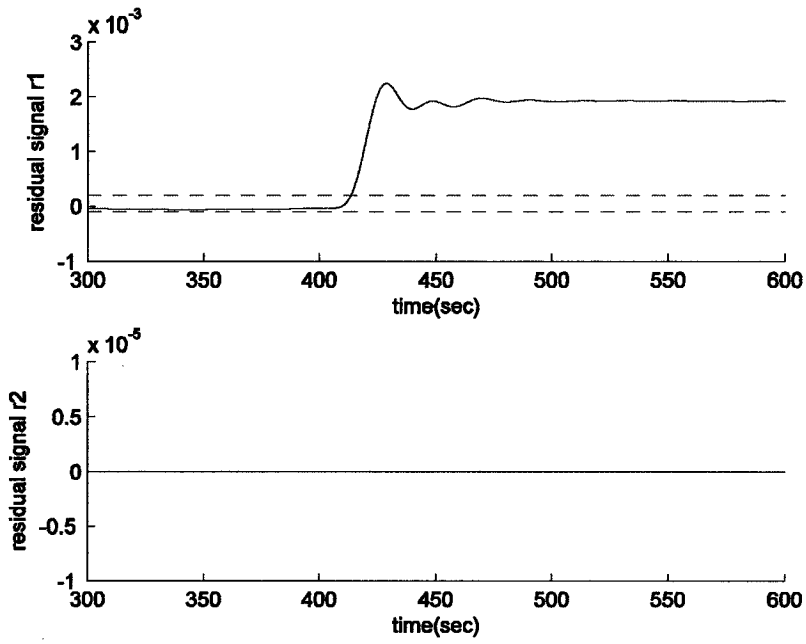


Figure 5-19: Detection of the fault at time 413sec by the residual signals generated due to a 70% fault in the gain k_t

5.2.1.2 Faults in the Viscous Friction Gain (τ_v)

The second source of fault that is considered is the gain τ_v. Figures 5-20, 5-21 and 5-22 depict the residual signals r_1 and r_2 for faults of different magnitudes in this gain. The time of occurrence of the faults in these simulations is 400sec and the faults are permanent. The fault sizes in these simulations are considered to be 20%, 50% and 80%. As can be seen in Figure 5-20, in some cases faults of as much as 20% cannot be detected. However Figures 5-21 and 5-22 show that faults of 50% and 80% can be detected by this observer. In these figures, after the fault occurs the residual signal crosses the lower boundary and consequently declares that a fault is present in the system. However, similar to the previous fault scenario in the motor torque gain, the

residual signal r_2 is not disturbed by the fault in the gain τ_v. Thus when a fault occurs, in the gain τ_v, it is expected that only the residual signal r_1 generates an alarm signal.

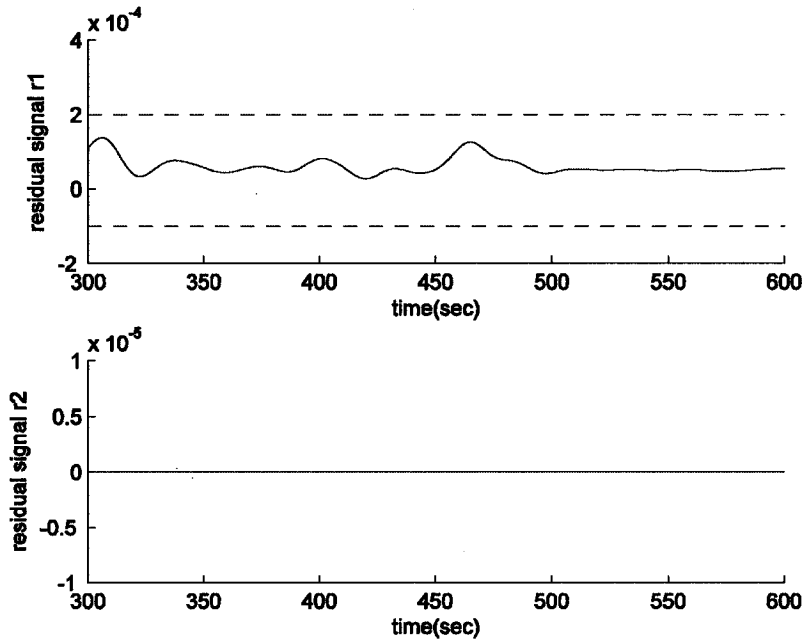


Figure 5-20: The residual signals due to a 20% fault in the gain τ_v

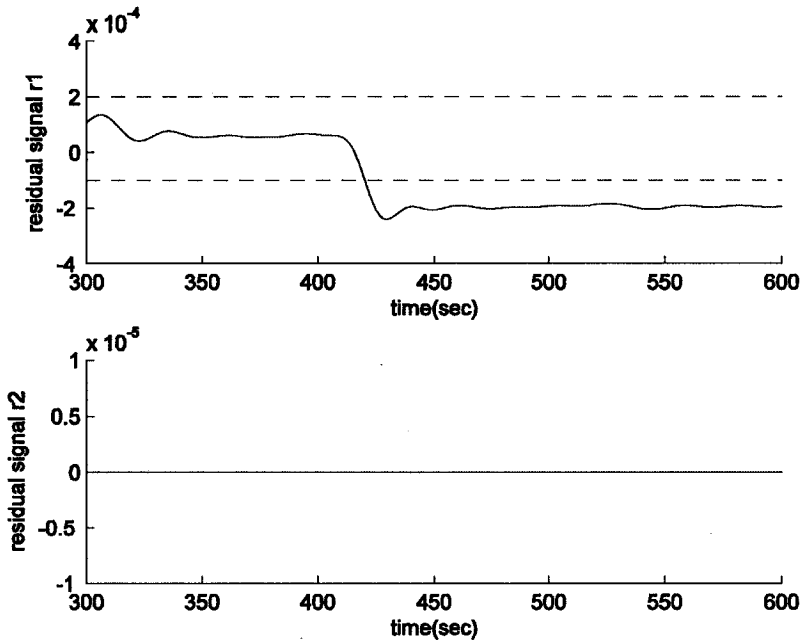


Figure 5-21: Detection of the fault at time 420sec by the residual signals generated due to a 50% fault in the gain τ_v

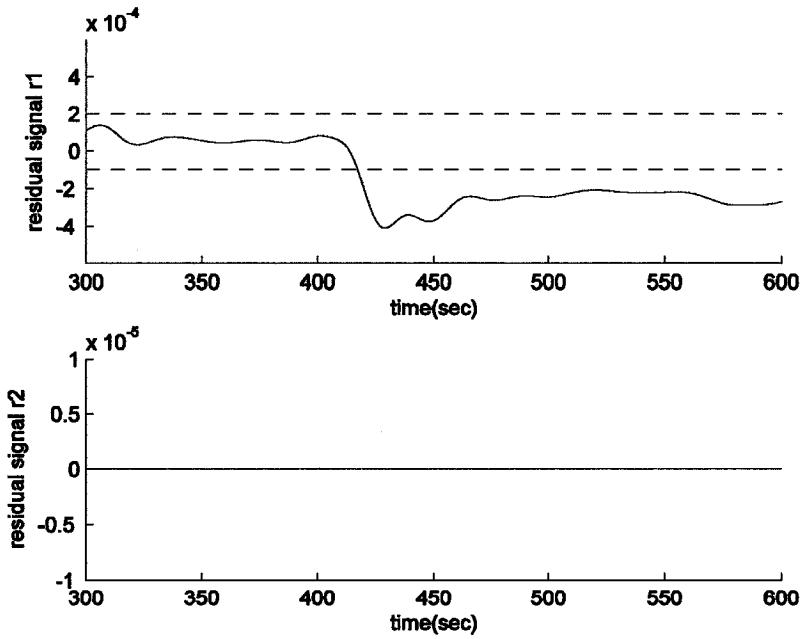


Figure 5-22: Detection of the fault at time 417sec by the residual signals generated due to an 80% fault in the gain τ_v

So far it could be noted that the residual signal r_1 is sensitive to both faults in the gains k_t and τ_v. That is, can detect both faults individually but it cannot identify the source of fault as it is excited by both. In other words, it cannot isolate these two types of fault from each other.

5.2.1.3 Faults in the Bus Voltage:

The last source of fault discussed is the bus voltage. In this section, the performance of the fault diagnosis algorithm confronting the fault in the bus voltage is tested. Figure 5-23, 5-24 and 5-25 demonstrate the residual signals when a fault happens in the bus voltage at 400sec and when it is permanent. The fault sizes are 30%, 60% and 90% respectively. In this scenario, it could be seen that both residuals are sensitive to the fault in the bus voltage. It was seen in the two previous fault scenarios that the residual signal r_2 is not sensitive to the faults in the gains k_t and τ_v. Therefore, the fault excitation of the residual signal r_2 can be considered to be equivalent to occurrence of fault in the bus voltage. This implies that the observer can identify the fault in the bus voltage, no matter if it is the only fault affecting the system or if there are multiple faults present in the system.

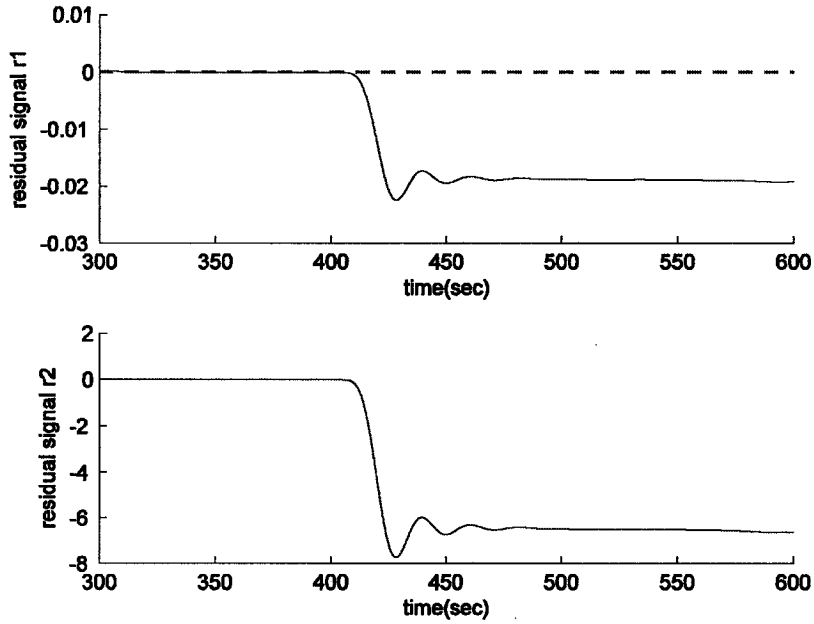


Figure 5-23: Detection of the fault at time 407sec by the residual signals generated due to a 30% fault in the bus voltage

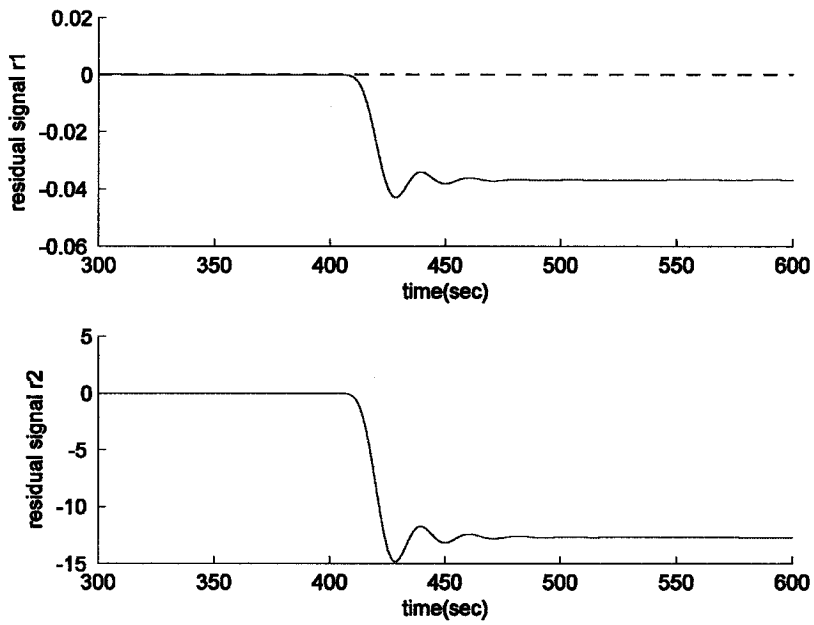


Figure 5-24: Detection of the fault at time 407sec by the residual signals generated due to a 50% fault in the bus voltage

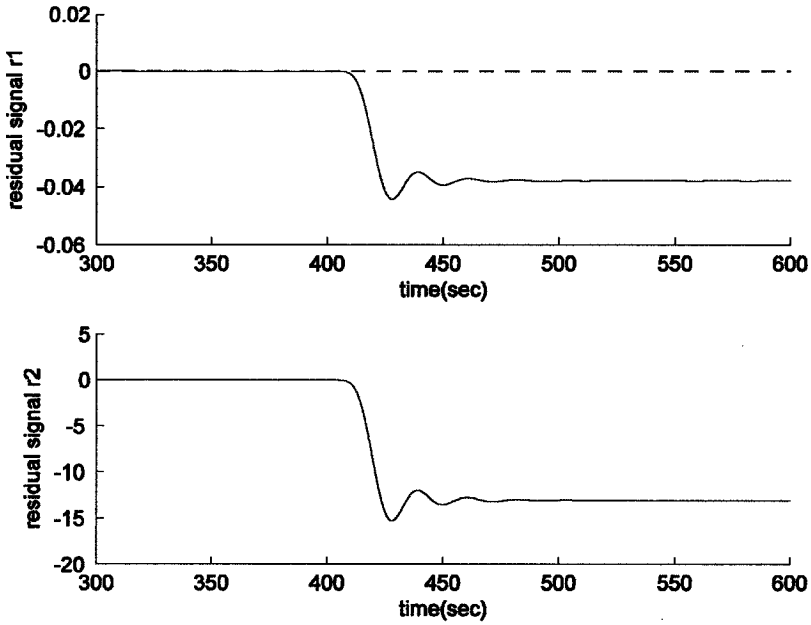


Figure 5-25: Detection of the fault at time 407sec by the residual signals generated due to a 70% fault in the bus voltage

5.2.1.4 Performance of the Fault Diagnosis Scheme in Presence of Intermittent Faults

A desired FDI algorithm is the one that maintains the alarming signal as long as the fault is present and persists in the system. In the last three sections, it was shown that the proposed FDI algorithm can achieve this objective. However, an alarm when there is no fault in the system is a false expression of the system state. A valid FDI scheme should stop all the warning signals when the faulty components return to their healthy mode. In this section, the performance of our proposed FDI scheme is investigated for the faults that are affecting the system in a finite duration of time.

As the first case, it is considered that there is a 40% fault in the gain k_t for a period of 100 seconds. Figure 5-26 shows this scenario. As expected, after a fault happens in the

gain k_t, the residual signal r_1 moves out of its no-fault zone and the signal r_2 remains unaffected. After the fault is removed from the system, it can be seen that the residual signal r_1 returns back to its no-fault interval while the residual signal r_2 is not excited again. Figure 5-27 shows a similar scenario for a fault in the gain τ_v. After 400sec the residual signal r_1, exceeds the safe band to identify the presence of a fault and it starts to return back to the safe band, 100 seconds after, at 500sec. The residual signal r_2, is unaffected during this time interval. Figure 5-28 demonstrates the behavior of the residual signals for a temporary fault in the bus voltage. Following the time that fault happens, both residuals are excited to show the presence of fault in the bus voltage and after the fault is no longer present in the system the residuals return to their normal operating condition.

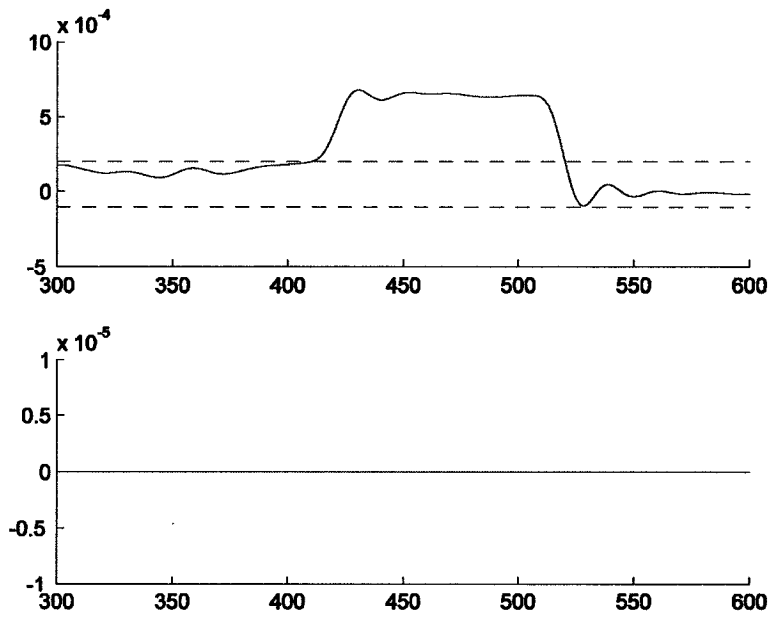


Figure 5-26: The residual signals due to an intermittent 40% fault in the gain k_t

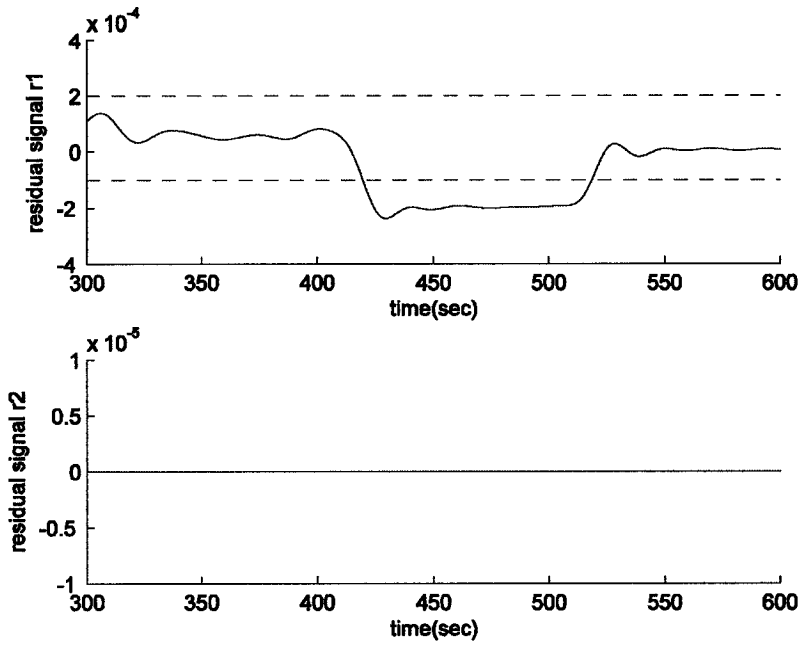


Figure 5-27: The residual signals due to an intermittent 50% fault in the gain τ_v

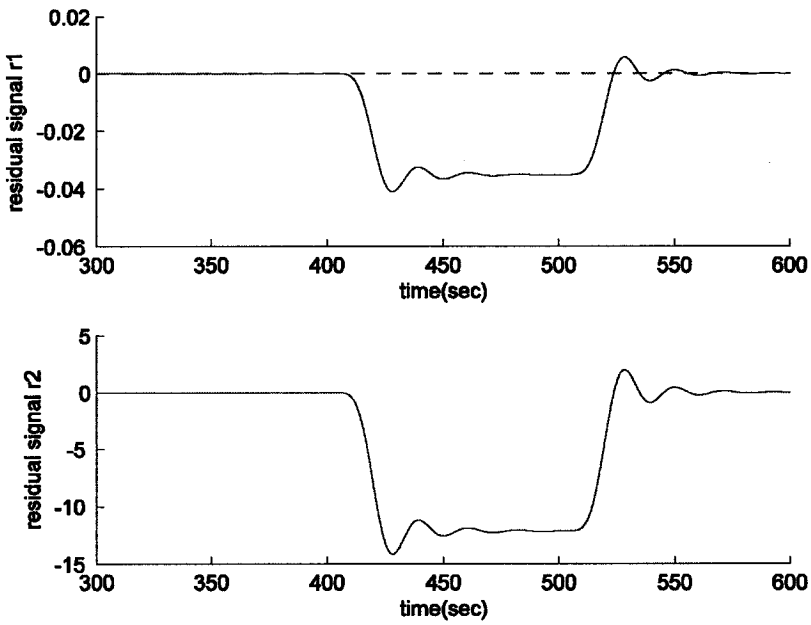


Figure 5-28: The residual signals due to an intermittent 60% fault in the bus voltage

5.2.1.5 Fault Isolation

So far the performance of the nonlinear observer as an FDI module has been investigated and studied for single faults. It was shown that all three types of faults can be detected using this fault diagnosis module. It was also shown that this module is capable of detecting the presence and absence of intermittent faults as well. In this section, the ability of this scheme is verified for isolation of faults. As was indicated in the previous section, residual signal r_1 is sensitive to all types of fault. On the other hand, the residual signal r_2 is only sensitive to the faults in the bus voltage. Therefore, any time that r_2 is triggered, it implies that a fault has occurred in the bus voltage. Whenever it is in its no-fault region, it is perceived that the bus voltage is not faulty. Consequently, when r_2 is not triggered and r_1 has exceeded one of its boundaries, it will be concluded that a fault in either gain k_t or the gain τ_v, or both has occurred. However, since these two types of fault are recognized through one residual signal, it is not possible to isolate them from one another using our proposed scheme without further processing of the signal r_1.

5.2.1.6 Performance of the Fault Diagnosis Algorithm in Presence of Concurrent Faults

All cases that have been discussed so far were dealing with scenarios in which the system is vulnerable to only one type of fault. In this section, the algorithm will be tested when multiple faults are present in the system. In order to evaluate the performance of the observer in presence of concurrent faults, different cases will be considered. In each case, the system is simulated when two sources of fault are affecting the system simultaneously.

(a) Presence of Concurrent Faults in the Gains k_t and τ_v

The first scenario that is considered is concurrency of faults in the gains k_t and τ_v. In the first case it is assumed that a 30% fault happens in the gain k_t and 100 seconds later an 80% fault happens in the gain τ_v. As can be seen from the Figure 5-29, the residual signal r_1 first exceeds the upper boundary to detect the occurrence of fault in the gain k_t. When a large 80% fault happens in the gain τ_v, it crosses through the no-fault band, passes the lower bound and remains outside the no-fault boundary thereafter. The residual signal r_2, on the other hand remains undisturbed during the entire time. Even when both faults are present after 500sec, this residual signal is not affected. In this case the warning system truly announces the presence of the fault from the beginning of the first fault and during the pursuit of the second fault. However, it cannot determine and specify the detailed information on how the system was affected by these faults.

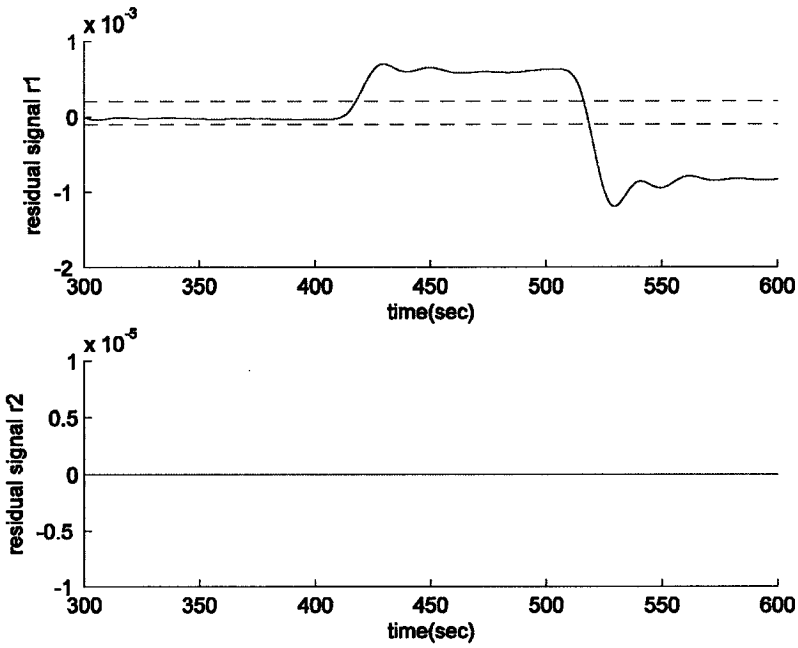


Figure 5-29: The residual signals due to concurrent faults in the gains k_t and τ_v

Note that in some cases a fault in viscous friction gain can cancel the effect of fault in motor torque gain and vice versa. Figure 5-30 shows that the residual signal r_1, detects a 30% fault in the gain k_t after its occurrence at 400sec. After a 30% fault happens in the gain τ_v at time 500sec, the residual signal returns to the no-fault zone. The residual signal r_2, on the other hand remains unaffected. The behavior of the signal r_1, after occurrence of fault in the gain τ_v, incorrectly indicates presence of no fault, while there are two faulty components in the system. In fact, the FDI scheme behaves as if there is a temporary fault in the gain k_t. Hence, this case shows that proposed nonlinear observer algorithm is not reliable when there are concurrent faults in the gains k_t and τ_v.

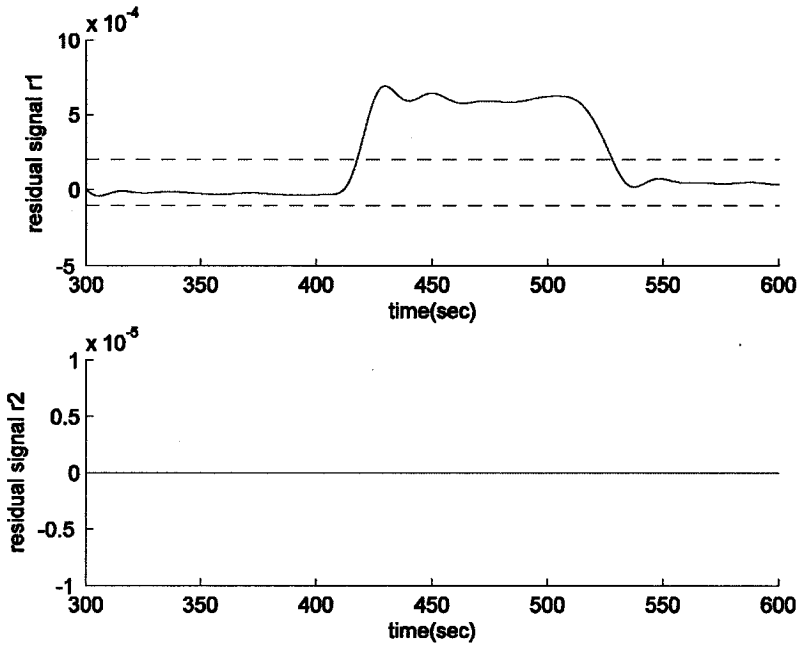


Figure 5-30: The residual signals due to concurrent faults in gains k_t and τ_v which leads to wrong diagnosis

(b) Presence of Concurrent Faults in the Gain k_t and the Bus Voltage

In this section, the behavior of the FDI algorithm is simulated when the gain k_t and the bus voltage are subject to faults. Figure 5-31 shows the residual signals generated by the observer when a 70% fault happens in the gain k_t at 400sec and a 30% fault happens in the bus voltage at 500sec. The observer first detects the fault in the gain k_t. This is achieved when the residual r_1 crosses its upper bound. When a fault happens in the bus voltage, both residuals are excited. If the fault in k_t is removed at some point, the observer cannot detect it. The reason is that due to the presence of a fault in the bus voltage, the signal r_1 remains far away from its no-fault zone. Therefore, the results from this simulation indicate that when a fault in the bus voltage occurs after a fault in the gain

k_t, it can be detected. Although after detection of the fault in the bus voltage, the proper detection of presence or absence of the fault in the gain k_t cannot be ensured any longer. In other words, if the fault in the motor torque gain is removed during the presence of fault in the bus voltage, it cannot be established.

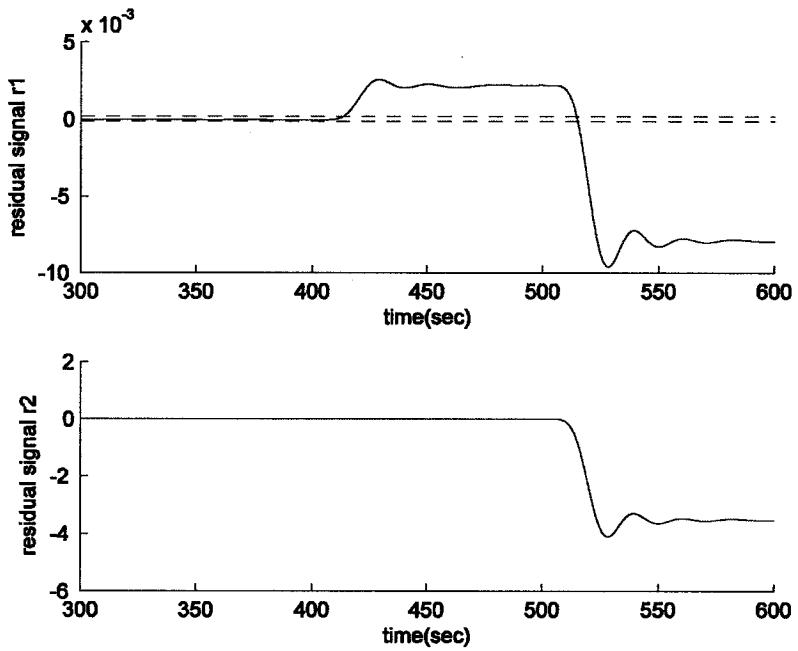


Figure 5-31: The residual signals due to concurrent faults in the gain k_t and the bus voltage

Following to the above scenario, let us assume that the fault in the bus voltage is removed at time 600sec, while the fault in the gain k_t is still present. Figure 5-32 demonstrates the residual signals for this scenario. As can be seen, after the time that the fault in the bus voltage is removed, the residual signal r_2 returns to zero while the residual signal r_1 remains outside its no-fault zone. Therefore, the residual generator correctly interprets this as disappearance of fault in the bus voltage and presence of fault in either gain k_t or τ_v.

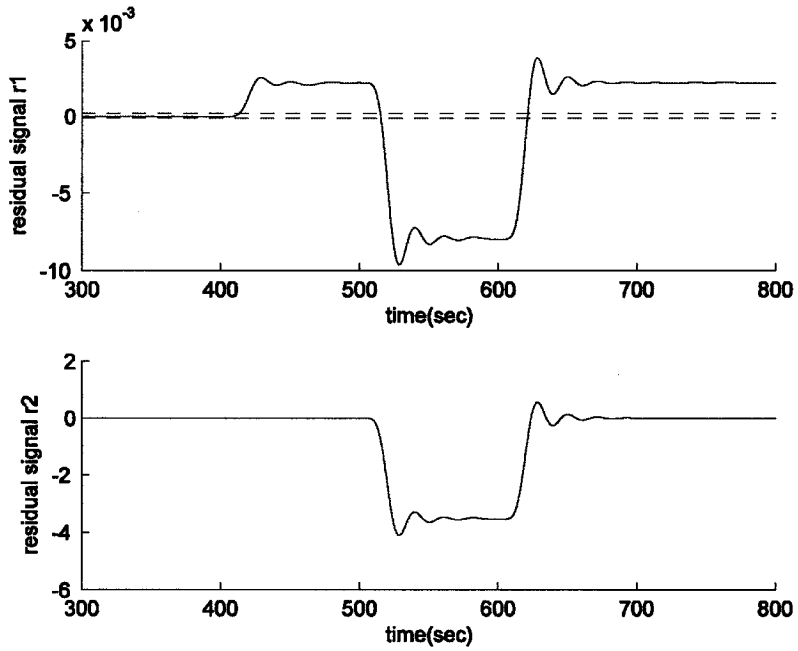


Figure 5-32: The residual signals due to concurrent faults in the gain k_t and the bus voltage

Figure 5-33, on the other hand, shows the residual signals when a fault happens in the bus voltage prior to a fault in the gain k_t. When a fault happens in the bus voltage both residual signals produce an alarm. This makes a motor torque gain fault insignificant to the observer. The reason is that this type of fault should be detected through divergence of the residual signal r_1 which is already in its faulty mode.

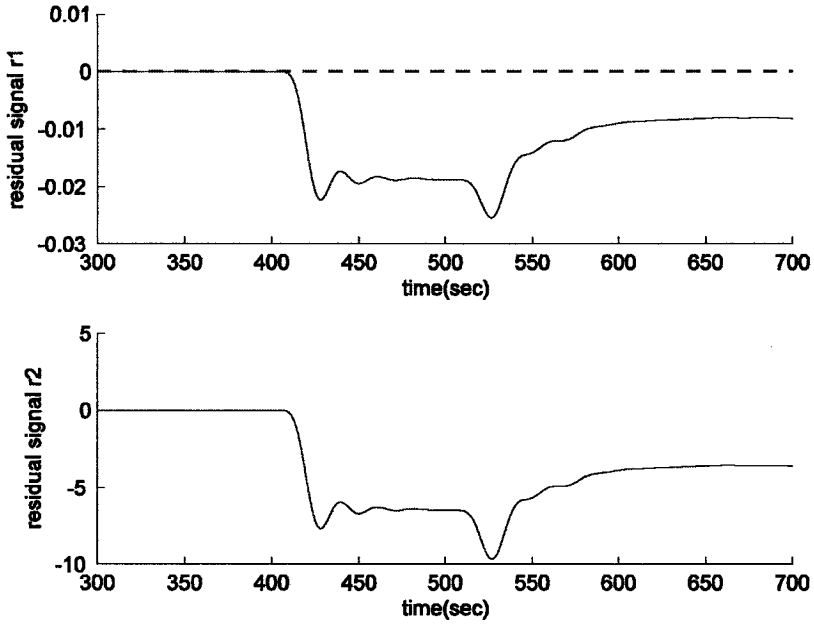


Figure 5-33: The residual signals due to concurrent faults in the bus voltage and the gain k_t

(c) Presence of Concurrent Faults in the Gain τ_v and the Bus Voltage

The last case of evaluating the effects of concurrent faults, involves the situation when the faults in the gain τ_v and bus voltage are considered. This case is very similar to the previous case, when the faults in the gain k_t and the bus voltage are overlapping. Figure 5-34 illustrates that the residual signal r_1 detects the fault in the gain τ_v. When a fault happens in the bus voltage after the occurrence of a fault in the gain τ_v, the residual signals remain in the faulty mode as long as the bus voltage fault is present. In this case the FDI scheme is not able to ensure if the fault in the gain τ_v is still disturbing the system. When the fault in the bus voltage happens before the one in the gain τ_v, a fault in gain τ_v can no longer be detected. This is shown in Figure 5-35.

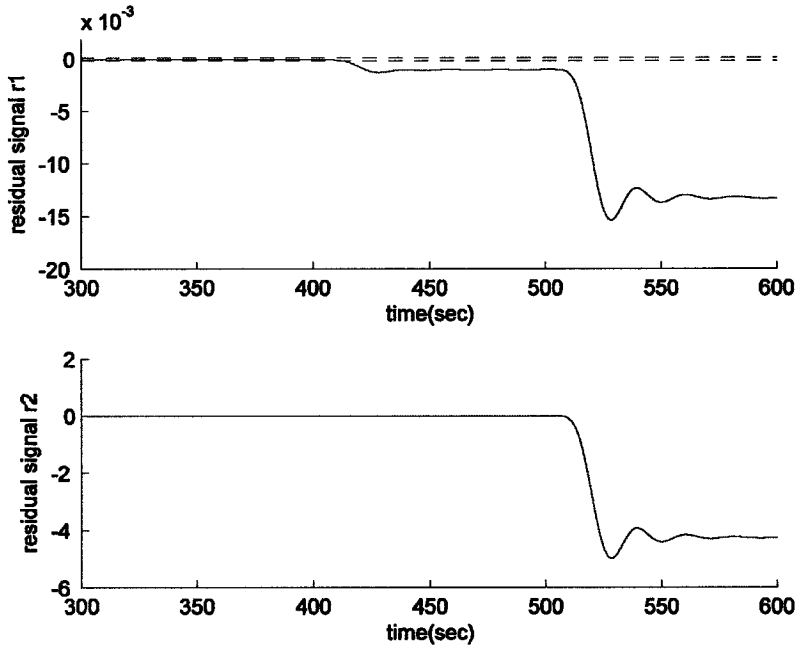


Figure 5-34: The residual signals due to concurrent faults in the gain τ_v and the bus voltage

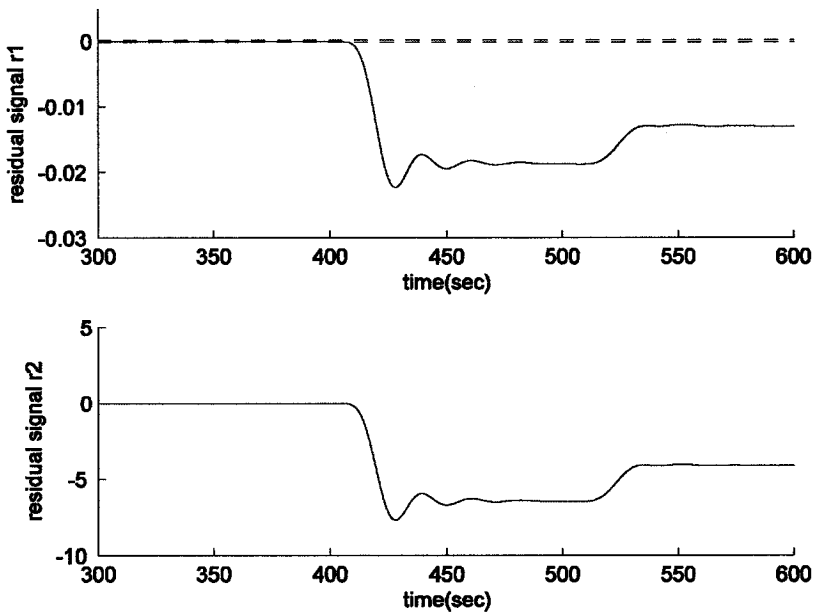


Figure 5-35: The residual signals due to concurrent faults in the bus voltage and the gain τ_v

5.2.2 Second Scenario: Low Bus Voltage, V_{bus}=6V

It was shown that a nonlinear observer can outperform the linear observer in the case of high bus voltages by detecting the faults in motor torque and viscous friction gains. Next the case of low bus voltage is considered.

5.2.2.1 Faults in the Motor Torque Gain (k_t)

Figures 5-36, 5-37 and 5-38 depict the diagnostic signals for faults of 10%, 40% and 70% in the gain k_t. It can be seen that after the fault happens at 800sec, both residual signals pass through their no-fault boundaries to detect the presence of a fault. Therefore, in low bus voltages, even very small faults in the motor torque gain can be detected.

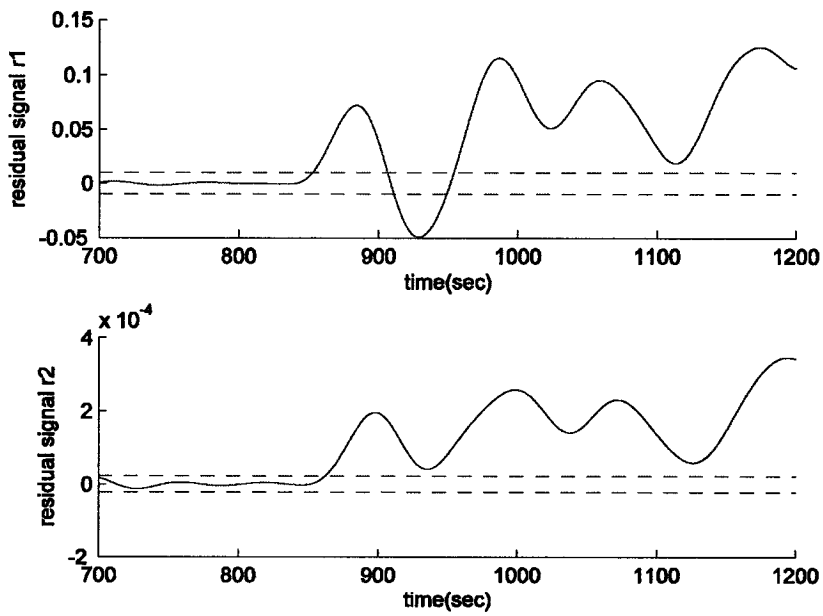


Figure 5-36: Detection of the fault at time 954sec by the residual signals generated due to a 10% fault in the gain k_t

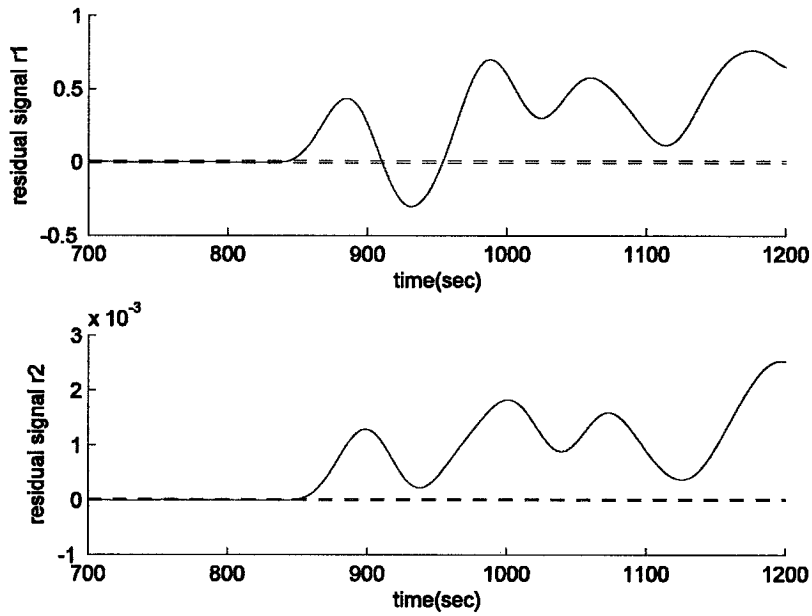


Figure 5-37: Detection of the fault at time 954sec by the residual signals generated due to a 40% fault in the gain k_t

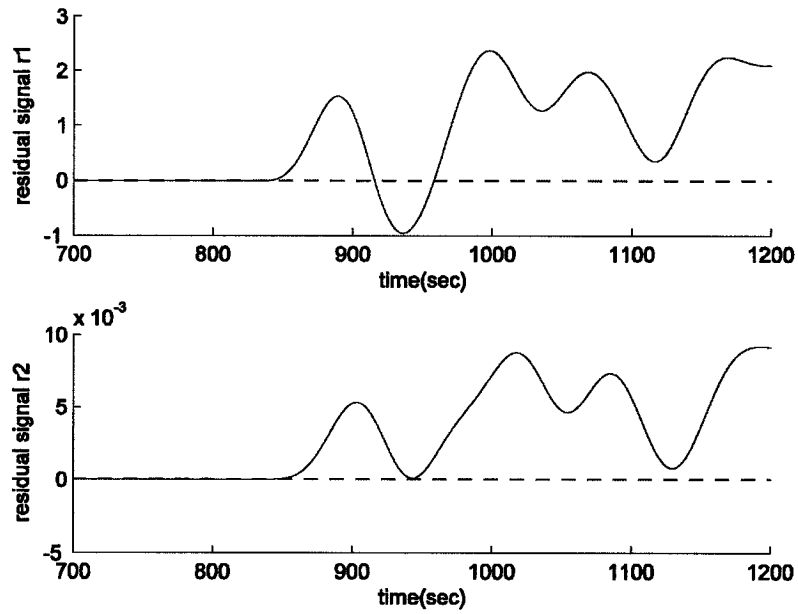


Figure 5-38: Detection of the fault at time 958sec by the residual signals generated due to a 70% fault in the gain k_t

5.2.2.2 Faults in the Viscous Friction Gain (τ_v)

Figures 5-39, 5-40 and 5-41 depict the diagnostic signals for faults of 20%, 50% and 80% in the gain τ_v. It can be seen that after the fault happens at 800sec, both residual signals are sensitive to this type of fault and the alarm signal is activated by both signals. Hence this type of fault can be detected using the nonlinear observer.

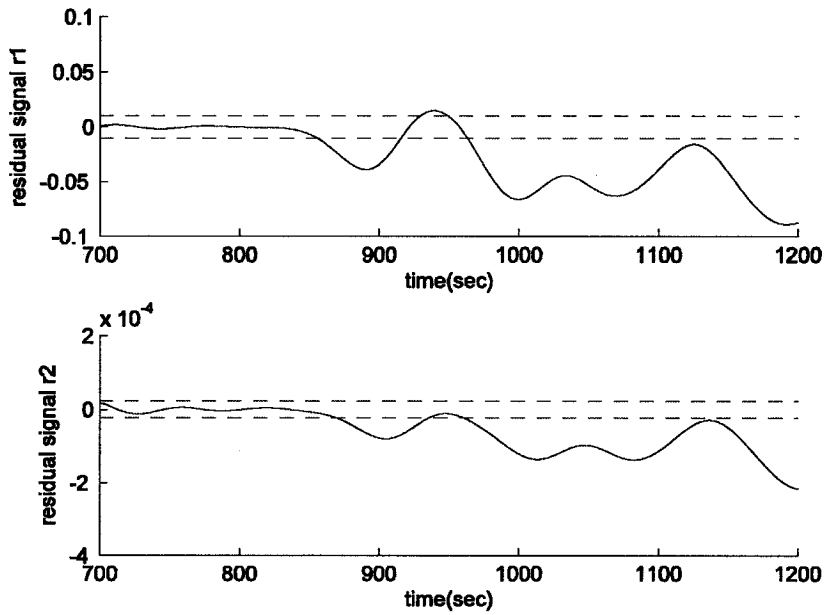


Figure 5-39: Detection of the fault at time 963sec by the residual signals generated due to a 20% fault in the gain τ_v

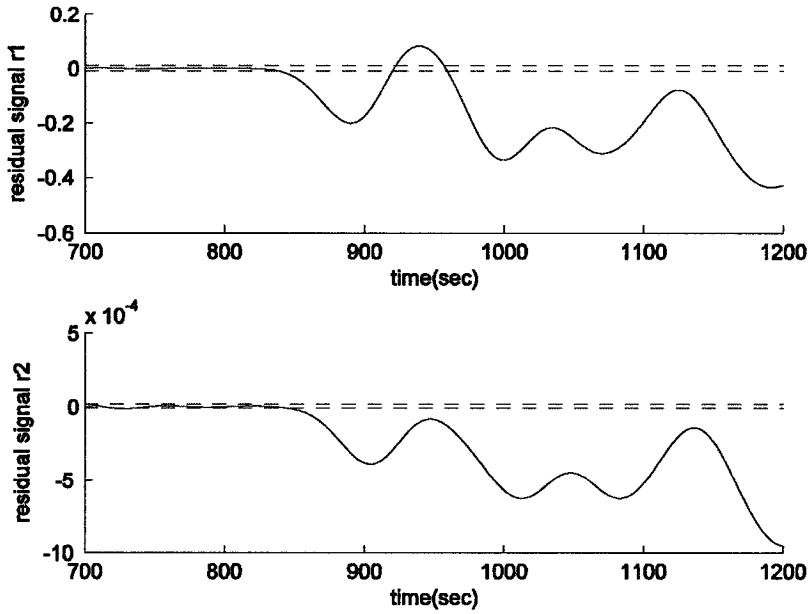


Figure 5-40: Detection of the fault at time 959sec by the residual signals generated due to a 50% fault in the gain τ_v

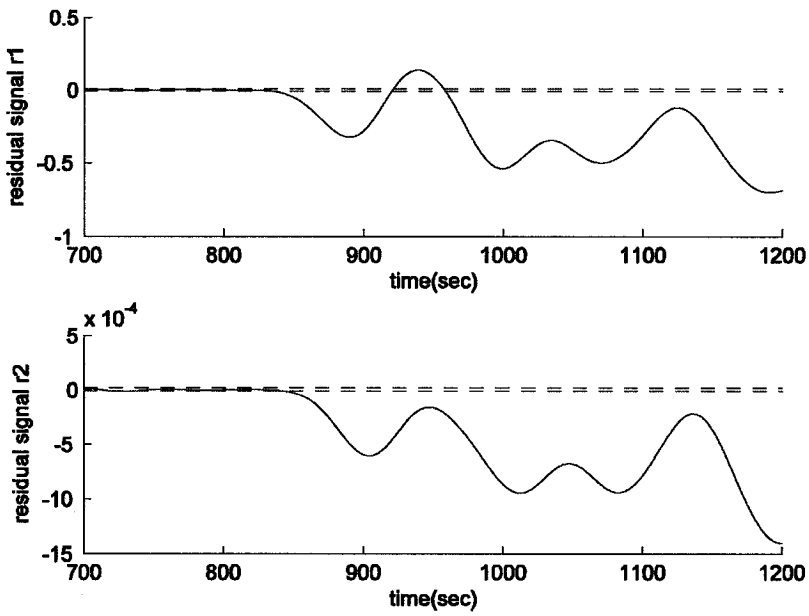


Figure 5-41: Detection of the fault at time 958sec by the residual signals generated due to an 80% fault in the gain τ_v

5.2.2.3 Faults in the Bus Voltage

Finally the performance of the nonlinear observer is investigated for faults in the bus voltage. Figures 5-42, 5-43 and 5-44 demonstrate the residual signals r_1 and r_2 when a fault takes place in the bus voltage at time 800sec. Similar to the case of high bus voltage, both residual signals are sensitive to this type of fault. However, in this case as was seen in the previous sections, the residual signal r_2 was sensitive to the faults in the gains k_t and τ_v as well. Therefore, there is no possibility of isolating this type of fault in the low bus voltage case.

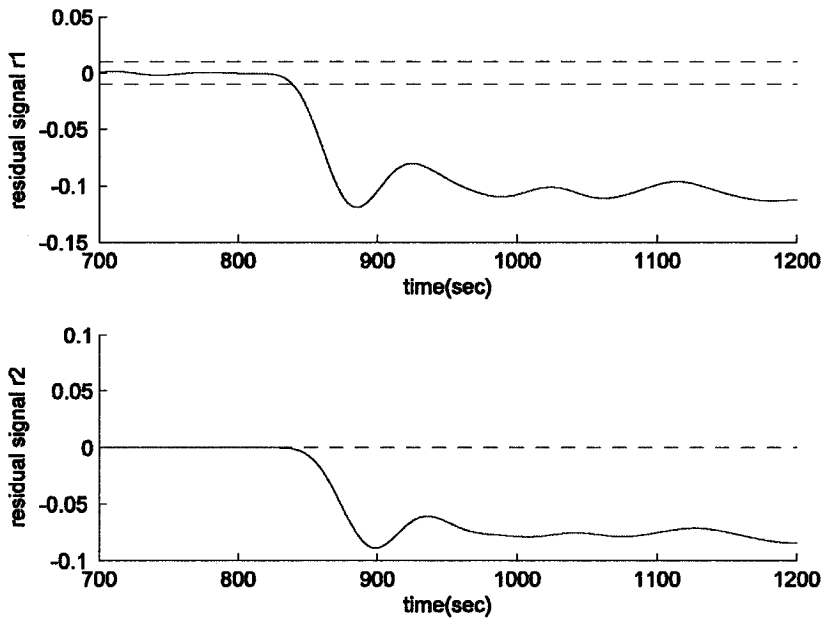


Figure 5-42: Detection of the fault at time 827sec by the residual signals generated due to a 10% fault in the bus voltage

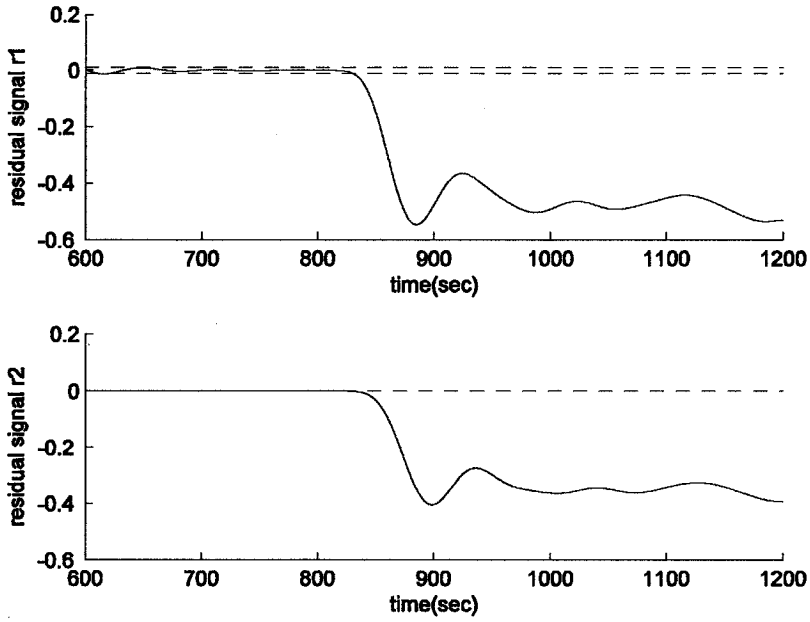


Figure 5-43: Detection of the fault at time 825sec by the residual signals generated due to a 40% fault in the bus voltage

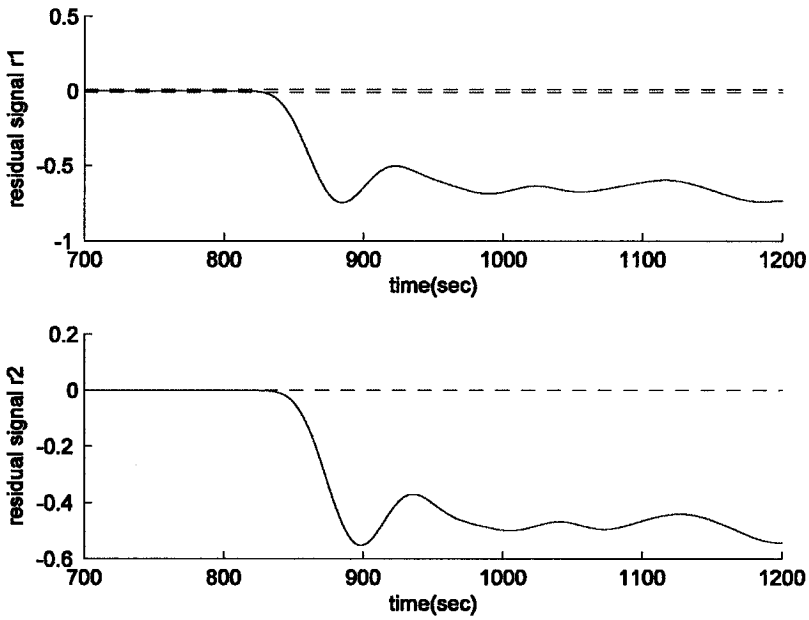


Figure 5-44: Detection of the fault at time 824sec by the residual signals generated due to a 50% fault in the bus voltage

5.2.2.4 Performance of the Fault Diagnosis Scheme in Presence of Intermittent Faults

Having inspected the performance of the nonlinear estimator for permanent faults, in this section the performance of the diagnosis module is investigated corresponding to intermittent faults. Figures 5-45, 5-46 and 5-47 demonstrate the residuals for intermittent faults all injected at time 800sec and removed at time 900sec. The residual signals all deviate from their no-fault regions when the fault is present and return to them when the fault is removed.

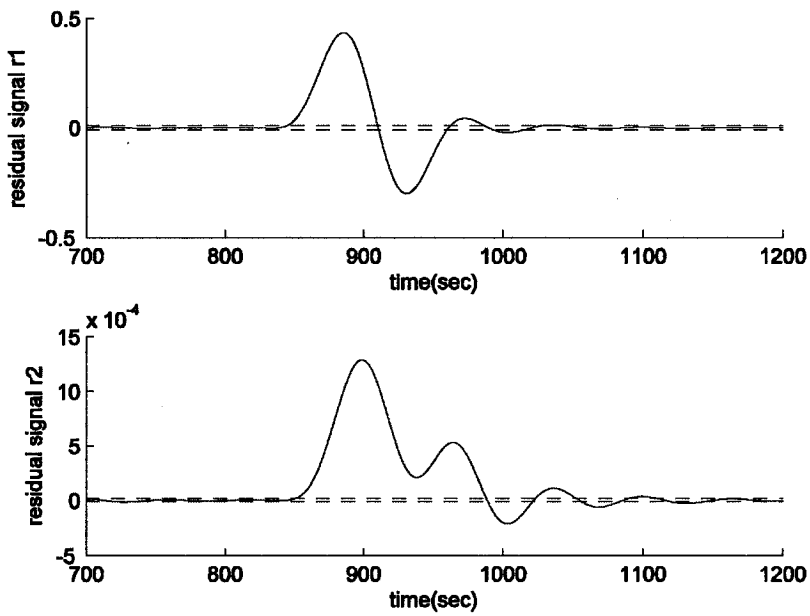


Figure 5-45: The residual signals due to an intermittent fault in the gain k_t

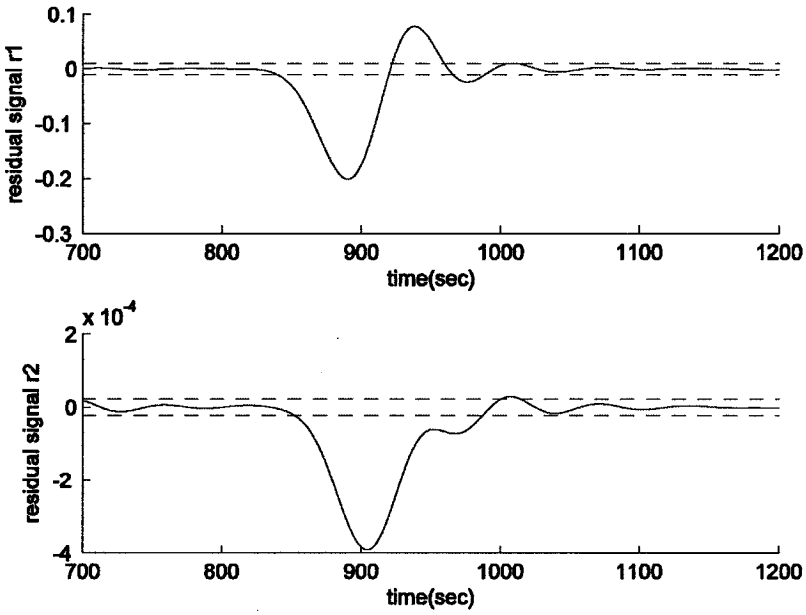


Figure 5-46: The residual signals due to an intermittent fault in the gain τ_v

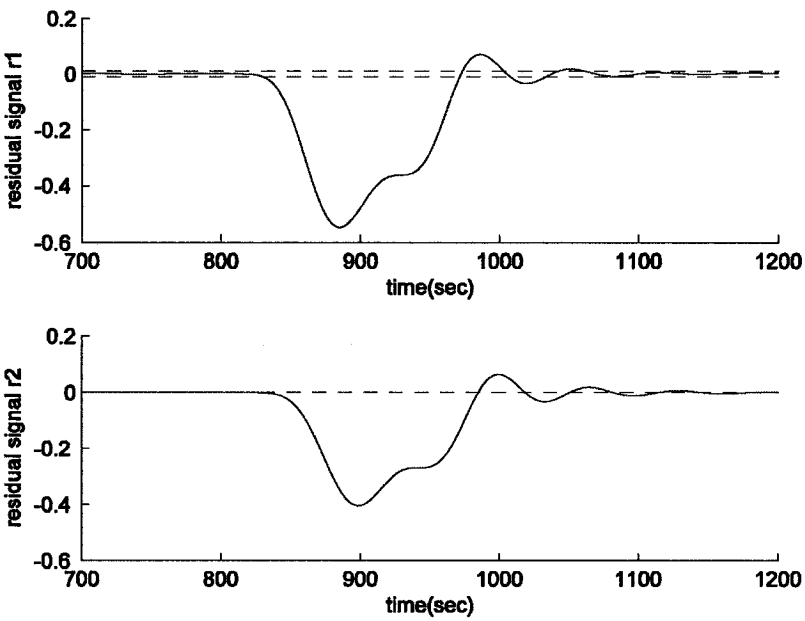


Figure 5-47: The residual signals due to an intermittent fault in the bus voltage

Since in the low bus voltage case both residuals are sensitive to all types of faults, investigating the performance of our proposed nonlinear observer for the presence of concurrent faults is trivial.

5.2.3 Presence of Fault in Reaction Wheels in Different Axes in Spacecraft Attitude Control System

In this section we investigate if the occurrence of a fault in the reaction wheel of one axis can have any effect on the residual signals corresponding to other axes. Since each diagnosis module in each axis is only supervising the reaction wheel associated to that axis, it is not expected that a fault in one reaction wheel influences a diagnosis module in another axis. However, different scenarios are simulated to show that the occurrence of faults in each reaction wheel does not disturb the residual signals corresponding to other reaction wheels. The faults in the following scenarios are considered to be very significant in order to show that even large faults cannot affect residual signals associated to other axes.

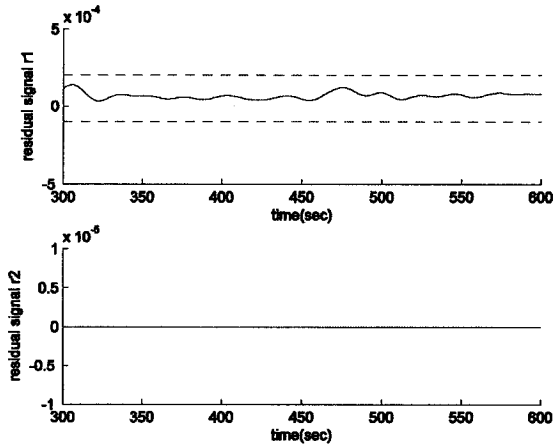
As the first scenario, it is assumed that an 80% fault occurs in the viscous friction gain in the reaction wheel associated to the yaw axis in the high bus voltage case. Figure 5-48 demonstrates the residual signals for the three axes in this scenario. As can be seen the residual signals associated to the yaw axis detect the presence of fault in the yaw axis. On the other hand, the residual signals of the other axes are not affected by this fault.

The second scenario is that of occurrence of a 70% fault in the bus voltage in the reaction wheel associated to the pitch axis in the high bus voltage case. Similar to the previous case Figure 5-49 shows that the diagnostic signals of the pitch axis detect this type of fault while those of the other axes remain in their no-fault zone.

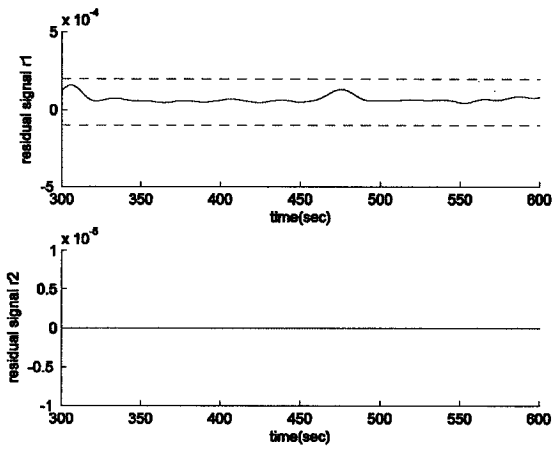
The third scenario deals with having a 70% fault in the motor torque gain of the reaction wheel in the roll axis in the low bus voltage case. Figure 5-50 illustrates the

residual signals for this scenario. As can be seen the residual signals of the roll axis warn about presence of a fault while the residual signals of the other axes are not disturbed.

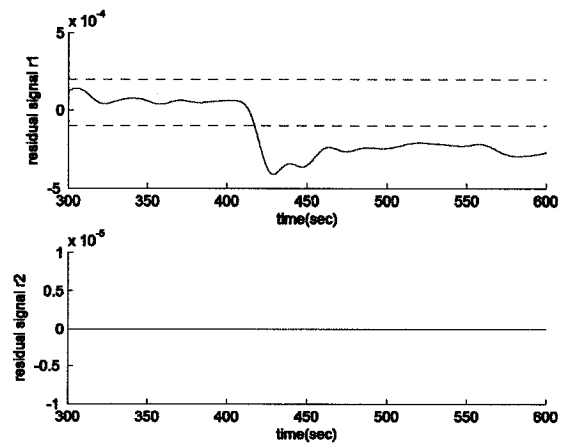
The last scenario to be considered is that of having a 50% fault in the bus voltage in the reaction wheel of the yaw axis in the low bus voltage case. Figure 5-51 demonstrates the residual signals for this experiment. It can be seen that the residual signals corresponding to the yaw axis detect the fault while the ones related to the other axes are not excited.



(a) The undisturbed residual signals associated to the roll axis

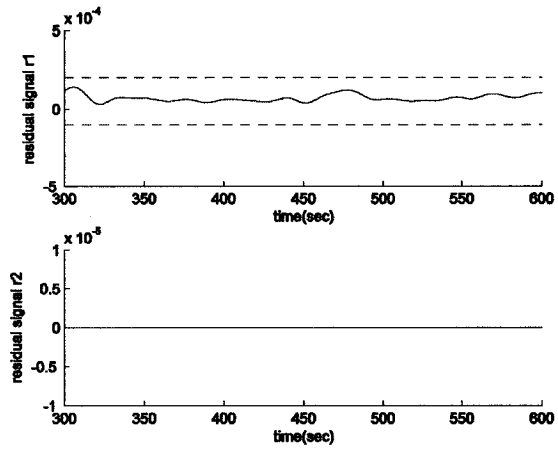


(b) The undisturbed residual signals associated to the pitch axis

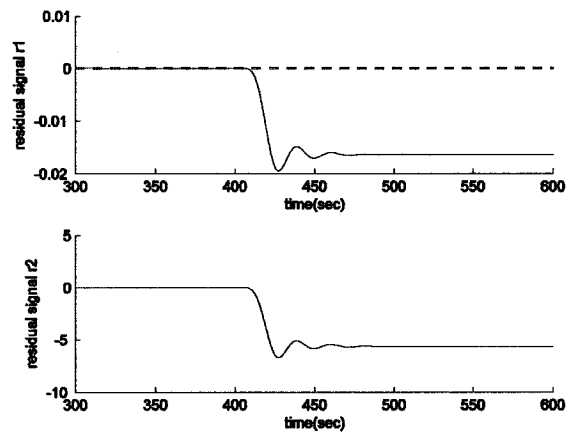


(c) Detection of an 80% fault in the gain τ_v by the residual signals associated to the yaw axis

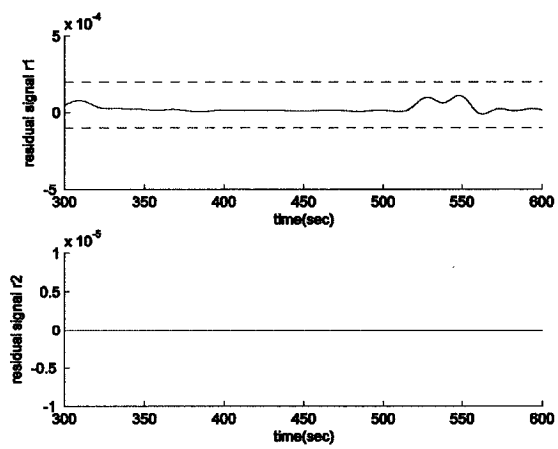
Figure 5-48: The three-axis residual signals for the high bus voltage case with fault in the yaw axis



(a) The undisturbed residual signals associated to the roll axis

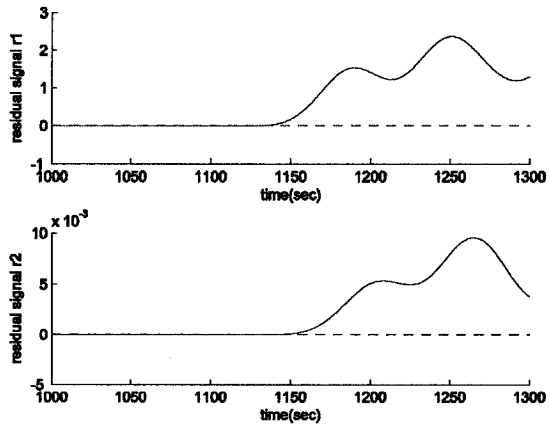


(b) Detection of a 70% fault in V_{bus} by the residual signals associated to the yaw axis

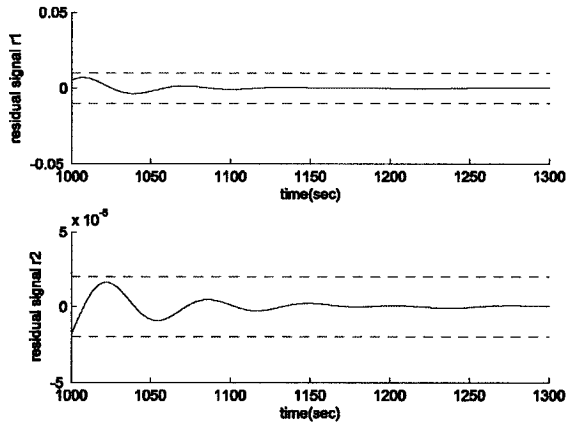


(c) The undisturbed residual signals associated to the yaw axis

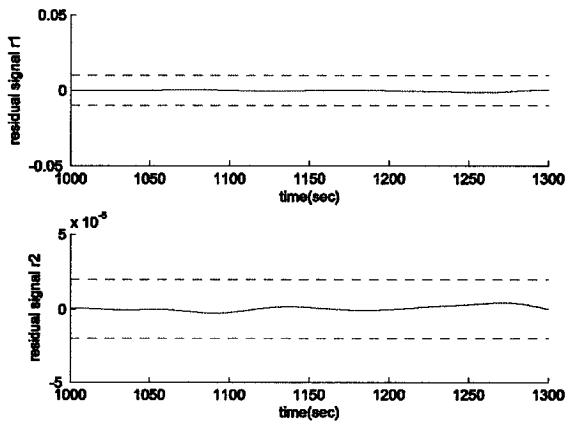
Figure 5-49: The three-axis residual signals for the high bus voltage case with fault in the pitch axis



(a) Detection of a 70% fault in the gain k_t by the residual signals associated to the roll axis

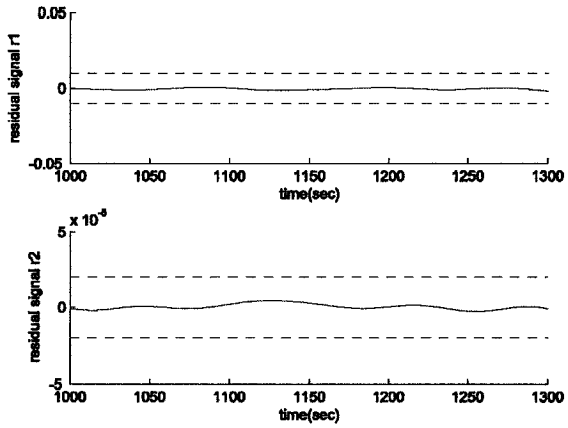


(b) The undisturbed residual signals associated to the pitch axis

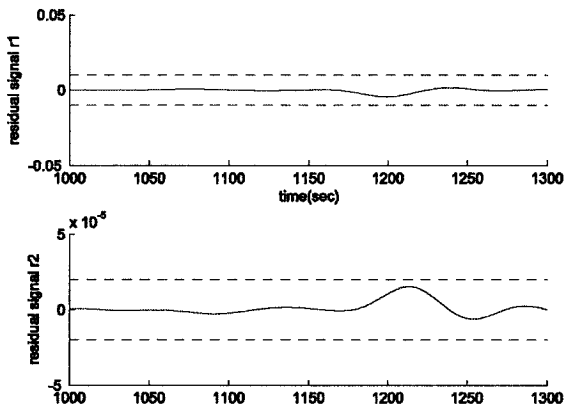


(c) The undisturbed residual signals associated to the yaw axis

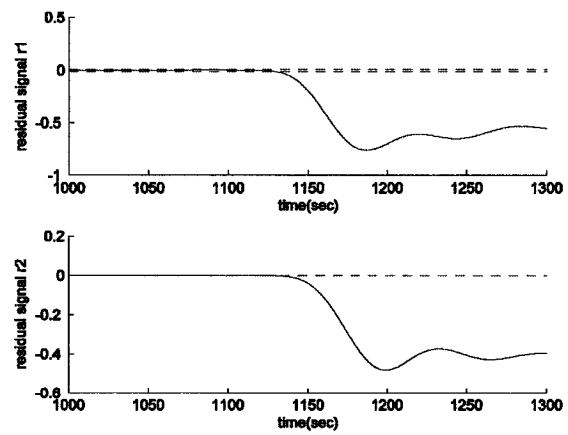
Figure 5-50: The three-axis residual signals for the low bus voltage case with fault in the roll axis



(a) The undisturbed residual signals associated to the roll axis



(b) The undisturbed residual signals associated to the pitch axis



(c) Detection of a 50% fault in V_{bus} by the residual signals associated to the yaw axis

Figure 5-51: The three-axis residual signals for the low bus voltage case with fault in the yaw axis

5.3 Discussion

In this chapter the performance of a linear observer versus the nonlinear observer was investigated and evaluated.

Table 5-1 shows the performance of a linear observer used as the FDI module. As was mentioned earlier this observer can only detect the faults in the bus voltage in the high bus voltage scenario. It does not detect faults in the motor torque and the viscous friction gains. In addition this linear observer is not convergent in low bus voltage scenario. The fault injection time, fault detection time and the delay in fault detection are indicated in this table.

Table 5-1: The performance of a linear observer for FDI

Bus Voltage (V)	Type of Fault	Fault Injection Time (sec)	Fault Detection Time (sec)	Delay in Fault Detection (sec)
12	10% fault in the motor torque gain	400	-	-
12	40% fault in the motor torque gain	400	-	-
12	70% fault in the motor torque gain	400	-	-
12	20% fault in the viscous friction gain	400	-	-
12	50% fault in the viscous friction gain	400	-	-
12	80% fault in the viscous friction gain	400	-	-
12	30% fault in the bus voltage	400	414	14
12	50% fault in the bus voltage	400	411	11
12	70% fault in the bus voltage	400	409	9

Table 5-2 shows the performance of a nonlinear observer used as the FDI module. This observer is able to detect faults in the motor torque gain, viscous friction gain and the bus voltage in both high bus voltage and low bus voltage scenarios. Similar to Table 5-1, the fault injection time, the fault detection time and the delay in fault detection are indicated in Table 5-2.

Table 5-2: The performance of a nonlinear observer for FDI

Bus Voltage (V)	Type of Fault	Fault Injection Time (sec)	Fault Detection Time (sec)	Delay in Fault Detection (sec)
12	10% fault in the motor torque gain	400	458	18
12	40% fault in the motor torque gain	400	413	13
12	70% fault in the motor torque gain	400	413	13
12	20% fault in the viscous friction gain	400	-	-
12	50% fault in the viscous friction gain	400	420	20
12	80% fault in the viscous friction gain	400	417	17
12	30% fault in the bus voltage	400	407	7
12	50% fault in the bus voltage	400	407	7
12	70% fault in the bus voltage	400	407	7
6	10% fault in the motor torque gain	800	954	154
6	40% fault in the motor torque gain	800	954	154
6	70% fault in the motor torque gain	800	958	158
6	20% fault in the viscous friction gain	800	855	155
6	50% fault in the viscous friction gain	800	845	145
6	80% fault in the viscous friction gain	800	843	143
6	10% fault in the bus voltage	800	827	27
6	40% fault in the bus voltage	800	825	25
6	50% fault in the bus voltage	800	824	24

Finally, Table 5-3 shows a comprehensive analysis and comparison between the two observers' performances. As can be seen the linear observer is successful in detecting the fault in only one fault scenario while the proposed nonlinear observer can detect the faults in all scenarios.

Table 5-3: The performance of a linear versus a nonlinear observer for FDI

Bus Voltage	Source of fault	Detected by the Linear observer	Detected by the Nonlinear observer
6v	Motor torque gain	×	✓
	Viscous friction gain	×	✓
	Bus voltage	×	✓
12v	Motor torque gain	×	✓
	Viscous friction gain	×	✓
	Bus voltage	✓	✓

5.4 Conclusion

In this chapter the performance of a linear observer as the fault diagnosis module was simulated. Since this type of observer was not successful in detecting all types of faults under all operating conditions of the system, the performance of a nonlinear observer designed in the previous chapter was investigated as an alternative. It was shown that this observer can outperform the linear observer and detect all types of faults under different operating conditions of the system.

Chapter 6

6 Conclusions and Future Work

In this thesis, an observer based approach was considered and analyzed as a fault diagnosis module in spacecraft attitude control system. A reaction wheel was considered as the actuator in the system. Reaction wheels are sensitive devices that are vulnerable to different sources of faults. The important role that is played by these wheels and the lack of any published research dealing with fault diagnosis of these devices have provided us with the motivation to develop a solution for this problem. A control law was first designed to stabilize the spacecraft system and to provide a framework for implementing the fault diagnosis algorithm.

First a linear observer was used as the fault diagnosis module. This observer could not asymptotically estimate the states of the system when the spacecraft was working with low bus voltages. In addition even in high bus voltages that it could yield proper state estimates, it could not detect the faults in the motor torque and viscous friction gains. The linear observer failure to solve the FDI problem has provided us with sufficient reasons to develop and obtain an alternative solution.

At the next step, a nonlinear observer was designed based on the parameters obtained from the linear observer. First, the state estimate convergence of this observer was

ensured analytically. Following the convergence verification, the proposed solution was applied as the fault diagnosis module in the nonlinear model of the spacecraft. It was assumed that the reaction wheel was prone to three different types of faults, namely faults in the motor torque gain, viscous friction gain and the bus voltage. The nonlinear observer generates two residual signals. Depending on the bus voltage operating condition, the performance of the nonlinear observer is investigated according to two cases.

In the first case a high bus voltage scenario was considered. In this case, one of the residual signals was only sensitive to one source of fault. On the other hand, the other residual signal was sensitive to all the three faults. Therefore, it was concluded that when there was only one type of a fault present in the system, it could be detected successfully. It was also shown that since one of the residual signals was only sensitive to faults in the bus voltage, this fault type could be isolated as well. Following the performance analysis of the proposed FDI scheme in presence of single faults, its performance was also tested for the concurrent fault scenarios. Given that both residual signals were excited by the presence of faults in the bus voltage, appearance or disappearance of a fault in the gains k_t and τ_v cannot be detected when there is a fault present in the bus voltage. When there is an overlap in the time of occurrence of faults in the gains k_t and τ_v, the FDI scheme can yield false information. As shown in Chapter 5, when one of these fault types is present in the system, occurrence of the other may be perceived as disappearance of the first fault. This is due to the fact that in this situation the residual signal returns to its normal operating mode when the second fault occurs.

In the second case a low bus voltage case was considered for examining the performance of the nonlinear observer. In this case both residual signals were susceptible to all types of faults. Therefore, all faults were detected but isolation was still not possible.

We can briefly list the contributions of this thesis as follows:

1. In this thesis we developed a fault diagnosis algorithm to diagnose the faults in a high-fidelity model of the reaction wheel in a fully nonlinear three-axis spacecraft attitude control system.
2. The developed algorithm can successfully detect all three types of fault in the reaction wheel. These three fault types are faults in the motor torque gain, viscous friction gain and the bus voltage.
3. This algorithm can isolate the faults in the bus voltage in the experiments dealing with high bus voltage, $V_{bus}=12V$.

Future Work

Future work can be directed towards achieving the following objectives:

1. **Isolation of all types of fault:** As shown in this thesis, the FDI scheme was not capable of isolating all faults. In an idealistic approach, three residual signals should be considered to isolate three sources of faults. Ideally each of these residuals should be affected by one and only one fault. In that case each fault will be associated to one residual signal. If such a residual generator is constructed, all types of fault can be clearly isolated. However this is not a necessary condition for isolating the faults. In our considered fault diagnosis

scheme two residual signals are generated. Since three sources of faults were defined in the FDI problem, the effects of at least two types of fault appear in one residual signal. Therefore, there are at least two fault sources that cannot be isolated. Future work may develop other signal processing methods to extract features that are uniquely identifiable to different faults. Development of these techniques which lead to isolating all sources of fault is worth investigation.

2. **Detection of concurrent faults:** Although for achieving the objective of isolating the faults there is no need to have a one to one matching between fault types and the residual signals, having such a relationship between residuals and faults can make the diagnosis valid even during concurrent faults. In this case the residual generator can distinguish all the existing faults from the others.

3. **Reducing the delay in fault detection:** Future work may attempt to develop techniques to decrease the delay in fault detection.

4. **Fault diagnosis in other components of the ACS:** Fault diagnosis schemes can also be developed for other components in the spacecraft attitude control system. This could be done for sensors (e.g. sun sensors, magnetometers, etc.) or other actuators such as electromagnets.

7 References

[1] E. Shokri, "Adaptive Fault-tolerance for Autonomous Spacecraft", Phase 1 Final Report, 1997.

[2] E. Shokri, "Adaptive Fault-tolerance for Autonomous Spacecraft", Phase 2 Final Report, 1999.

[3] J. Chen, R. J. Patton, Robust Model-Based Fault Diagnosis for Dynamic Systems, Boston, 1999.

[4] R. V. Beard, "Failure accommodation in linear systems through self-re-organization," Ph.D. dissertation, Mass. Inst. Technol., Cambridge, MA, 1971.

[5] H. L. Jones, "Failure detection in linear systems," Ph.D. dissertation, Mass. Inst. Technol., Cambridge, MA, 1973.

[6] M. –A. Massoumnia, G. C. Verghese, and A. S. Willsky, "Failure detection and identification," IEEE Trans. Automat. Contr., vol. AC-31, pp. 839-846, 1986.

[7] V. Venkatasubramanian, R. Rengaswami, S. N. Kavuri, K. Yin, "A Review of Fault Detection and Diagnosis, Part III: Process History Based Methods", Elsevier Press, Computers and Chemical Engineering 27, pp. 327-346, 2003.

[8] S. A. Lapp, and G. A. Powers, "Computer-Aided synthesis of fault trees", IEEE Trans. Reliability, 37, 2-13, 1977.

[9] P. M. Frank, "Fault Diagnosis in dynamic systems using analytical and knowledge-based redundancy-A survey," Automatica, vol. 26, pp. 459-474, 1990.

[10] A. J. Krener and A. Isidori, "Linearization by output injection and nonlinear observers," Syst. Control Lett., vol. 3, pp. 47-52, 1983.

[11] G. Besancon and H. Hammouri, "On uniform observation of nonuniformly observable systems," Sys. Control Lett., vol. 29, pp. 9-19, 1996.

[12] L. Praly and Z. P. Jiang, " Stabilization by output feedback for systems with ISS inverse dynamics," Syst. Control Lett., vol. 21, pp. 19-33, 1993.

[13] H. K. Khalil and F. Esfandiari, "Semiglobal stabilization of a class of nonlinear systems using output feedback," IEEE Trans. Automat. *Contr.*, vol. 38, pp. 1412-1415, Sept. 1993.

[14] A. Teel, L. Praly, "Global stabilizability and imply semi-global stabilizability by output feedback," Syst. Control Lett., vol. 22, pp. 313-325, 1994.

[15] L. A. Mironovskii, "Functional diagnosis of linear dynamic systems," Automat. Remote Contr., pp. 1198-1205, 1980.

[16] E. Y. Chow and A. S. Willsky, "Analytical redundancy and the design of robust failure detection systems," IEEE Trans. Automat. Contr., vol. AC-29, pp. 603-614, 1984.

[17] J. J. Gertler, K. C. Anderson, An evidential reasoning extension to quantitative model-based failure diagnosis Systems, Man and Cybernetics, IEEE Transactions on Volume 22, Issue 2, pp. 275 – 289, 1992

[18] E. J. Henley, Application of expert systems to fault diagnosis In AIChE annual meeting, San Francisco, CA, 1984.

[19] D. Chester, D. Lamb, P. Dhurjati, Rule-based computer alarm analysis in chemical process plants. In Proceedings of 7th Micro-Delcon. pp. 22-29, 1984.

[20] V. Venkatasubramanian and S. H. Rich, An object-oriented two-tier architecture for integrating compiled and deep-level knowledge for process diagnosis. Computers and Chemical Engineering 12 (9-10), pp. 903-921, 1988.

[21] T. S. Ramesh, J. F. Davis, G. M. Schwenzer, Catcracker: an expert system for process and malfunction diagnosis in fluid catalytic cracking units. In AIChE annual meeting, San Francisco, CA, 1989.

[22] M. Basila, Jr., G. Stefanek, and A. Cinar, A model-object based supervisory expert system for fault tolerant chemical reactor control. Computers and Chemical Engineering 14 (4- 5), pp. 551- 560, 1990.

[23] J. T. Cheung, and G. Stephanopoulos, Representation of process trends part I. A formal representation framework. Computers and Chemical Engineering 14 (4-5), pp. 495-510, 1990.

[24] M. Janusz, and V. Venkatasubramanian, Automatic generation of qualitative description of process trends for fault detection and diagnosis. Engineering Applications of Artificial Intelligence 4 (5), pp. 329-339, 1991

[25] M. Basseville, and T. V. Nikiforov, Detection of abrupt changes-theory and application. Information and system sciences series. Prentice Hall, 1993.

[26] W. A. Shewhart, Economic control of quality of manufactured product. Princeton, NJ: Van Nostrand, 1931.

[27] K. Pearson, On lines and planes of closest fit to systems of points in space. Philosophical Magazine Series B 2, pp. 559-572, 1901

[28] T. W. Anderson, An introduction to multivariate statistical analysis, New York: Wiley, 1984.

[29] J. E. Jackson, A user's guide to principal components. New York: Wiley-Interscience, 1991.

[30] S. Wold, Cross-validatory estimation of the number of components in factor and principal components models. Technometrics 20 (4), pp. 397-405, 1978.

[31] S. Wold, K. Esbensen, and P. Geladi, Principal component analysis. Chemometrics and Intelligent Laboratory Systems 2 (1-3), pp. 37-52, 1987.

[32] H. Wold, Soft modeling, the basic design and some extensions. In K. Joreskog & H. Wold (Eds.), System under indirect observations. Amsterdam: North Holland, 1982.

[33] J. F. Macgregor, J. Christina, K. Costas and M. Katoudi, "Process Monitoring and Diagnosis by multi-block PLS methods", AIChE J, 40(5), pp. 826-838, 1994.

[34] L. H. Ungar, B. A. Powell, and S. N. Kamens, Adaptive networks for fault diagnosis and process control. Computers and Chemical Engineering 14 (4-5), pp. 561-572, 1990.

[35] J. C. Hoskins, K. M. Kaliyur, and D. M. Himmelblau, Fault diagnosis in complex chemical plants using artificial neural networks. American Institute of Chemical Engineers Journal 37 (1), pp. 137-141, 1991.

[36] J. Y. Fan, M. Nikolaou, and R. E. White, An approach to fault diagnosis of chemical processes via neural networks. American Institute of Chemical Engineers Journal 39 (1), pp. 82-88, 1993.

[37] A. E. Farell, and S. D. Roat, Framework for enhancing fault diagnosis capabilities of artificial neural networks. Computers and Chemical Engineering 18 (7), pp. 613-635, 1994.

[38] C. S. Tsai, and C. T. Chang, Dynamic process diagnosis via integrated neural networks. Computers and Chemical Engineering 19, pp. 747-752, 1995.

[39] W. Becraft, and P. Lee, An integrated neural network/expert system approach for fault diagnosis. Computers and Chemical Engineering 17 (10), pp. 1001-1014, 1993.

[40] M. D. Montemarlo, "The AI program at the National Aeronautics and Space Administration", AI magazine, pp. 49-61, 1992.

[41] M. M. Polycarpou, and M. A. Helmicki, "Automated fault detection and accommodation", IEEE Trans. Syst., Man, Cybern., vol. 25, pp. 1447-1458, 1995.

[42] H. Wang, Z. J. Huang, and S. Daley, "On the use of adaptive updating rules for actuator and sensor fault diagnosis," Automatica, vol. 33, pp. 217-225, 1997.

[43] A. T. Vemuri, and M. M. Polycarpou, "On-line approximation-based methods for robust fault detection," in Proc. 13th IFAC World Congr., pp. 319-324, 1996.

[44] J. Gertler, Fault detection and diagnosis in engineering systems, Marcel Dekker, 1998.

[45] N. Kuipel and P. M. Frank, A fuzzy FDI Decision-Making System for the Support of the Human Operator, In Patton and Chen(eds), Proc. IFAC Symposium SAFEPROCESS'97, Hull, UK, 721-726, Pergamon Press, 1998.

[46] R. J. Patton, C. J. Lopez-Toribio, Artificial intelligence approaches to fault diagnosis, Update on Developments in Intelligent Control (Ref. No. 1998/513), IEE Colloquium, pp. 3/1 – 3/12, 1998.

[47] W. S. Lee, D. I. Grosh, Tillman F.A. and Lie, C.H., "Fault tree analysis, methods and applications- A review," IEEE Transactions on Reliability, vol. R-34, no. 3, pp. 194-203, 1985.

[48] L. E. Holloway, and S. Chand, "Time templates for discrete event fault monitoring in manufacturing systems," Proc. Amer. Contr. Conf., Baltimore, MD, pp. 701-706, 1994.

[49] F. Lin, "Diagnosability of discrete event systems and its applications," Discrete Event Dynamic Systems, vol.4, pp. 197-212, 1994.

[50] A. Bouloutas, G. Hart, and M. Shwartz, "On the design of observers for fault detection in communication networks," in Network Management and Control, A. Kershenbaum, M. Malek and M. Wall, Eds, New York: Plenum Press, 1990.

[51] M. Sampath, R. Sengupta, S. Lafortune, K. Sinnamohideen, and D. Teneketzis, "Diagnosability of discrete-event systems," IEEE Transactions on Automatic Control, vol-40, no.9, pp. 1555-1575, 1995.

[52] M. Sampath, R. Sengupta, S. Lafortune, K. Sinnamohideen, and D. Teneketzis, "Failure diagnosis using discrete-event models," IEEE Transactions on Control Systems Technology, vol.4, no.2, pp. 105-124, 1996.

[53] D. N. Pandalai, and L. E. Holloway, "Template languages for fault monitoring of timed discrete-event processes," IEEE Transactions on Automatic Control, vol.45, no.5, pp. 868-882.

[54] A. S. Korovkin, Spacecraft Control Systems, NASA technical translation, Washington, D. C., 1973.

[55] T. L. Wilson, "A high-temperature superconductor energy-momentum control system for small satellites" Applied Superconductivity, IEEE Transactions on Volume 13, Issue 2, pp. 2287 – 2290, 2003.

[56] R. Froelich, H. Papapoff, "Reaction wheel attitude control for space vehicles" Automatic Control, IRE Transactions on, Volume: 4, Issue: 3, pp. 139 – 149. 1959.

[57] J. R. Wertz and W. J. Larson, Space Mission Analysis and Design, Third Edition, Kluwer Academic Publishers, 1999.

[58] B. Bialke, "High fidelity mathematical modeling of reaction wheel performance," in 21st Annual American Astronautical Society Guidance and Control Conference, AAS paper 98-063, 1998.