Failure Detection and Isolation for the Space Shuttle Main Engine

by

Nhut Tan Ho

Submitted to the Department of Mechanical Engineering
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN MECHANICAL ENGINEERING

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1998   [June 1998]

Author...

Department of Mechanical Engineering
May 20, 1998

Certified by ....
Dr. Rami S. Mangoubi
Senior Member of the Technical Staff, C.S. Draper Laboratory
Thesis Supervisor

Certified by ....
Mr. Roger M. Hain
Senior Member of the Technical Staff, C.S. Draper Laboratory
Thesis Supervisor

Certified by ....
Professor Derek Rowell
Professor of Mechanical Engineering
Thesis Supervisor

Accepted by ....
Dr. Ain Sonin
Chairman, Department Graduate Committee

Failure Detection and Isolation for the Space Shuttle Main Engine

by

Nhut Tan Ho

Submitted to the Department of Mechanical Engineering
on May 20, 1998, in partial fulfillment of the
requirements for the degree of
MASTER OF SCIENCE IN MECHANICAL ENGINEERING

Abstract

In this thesis, a Vehicle Health Monitoring (VHM) system is designed to detect and isolate failures in the engines of Reusable Launch Vehicles (RLV). In particular, our focus has been on the Space Shuttle Main Engine (SSME). This VHM system is an improvement over the current safety monitoring system that relies on redline methods to detect failures. This VHM system takes into account engine failures in both sensors and valves, as well as internal components such as turbopumps, injectors, and the combustion chamber. The algorithm of the VHM system is model based. Specifically, using information from a thermodynamic model of the engine together with sensor measurements, a Kalman filter (KF) is designed to predict sensor outputs. The residual, or the difference between the predicted and actual measurement, is used by several statistical tests to detect the presence of a failure, and to categorize the failure as a sensor/valve failure or an internal component failure. Sensor and valve failures are diagnosed using the Generalized Likelihood Ratio Test (GLRT). Internal component failures, on the other hand, are isolated using the Multiple Model (MM) method. The proposed methodology can be used for online Failure Detection and Isolation (FDI) as well as for postflight analysis. At the design stage, it can help determine the detectability and distinguishability of failures given a candidate sensor configuration.

Thesis Supervisor: Dr. Rami S. Mangoubi
Title: Senior Member of the Technical Staff, C.S. Draper Laboratory

Thesis Supervisor: Mr. Roger M. Hain
Title: Senior Member of the Technical Staff, C.S. Draper Laboratory

Thesis Supervisor: Professor Derek Rowell
Title: Professor of Mechanical Engineering

Acknowledgments

Publication of this thesis does not constitute approval by the Draper Laboratory or the sponsoring agency of the findings or conclusions contained therein. It is published for the exchange and stimulation of ideas.

Permission is hereby granted by the Author to the Massachusetts Institute of Technology to reproduce any or all of this thesis.

Contents

List of Figures

9

List of Tables

Chapter 1

Introduction

The increasing demand for conducting experiments in space, and installing satellites into orbits for commercial facilities, have necessitated the use of Reusable Launch Vehicles (RLV) more than ever before. In essence, this demand motivates the development and design of high performance, sophisticated technologies that significantly improve the performance of the RLV. These technologies would not only aim to reduce vehicle life cycle cost and decrease fuel consumption but also encompass control system that counteract problems such as rejection of external disturbance in uncertain environments, and detection and compensation of system failures.

This thesis addresses the problem of failure detection for the RLV's. Specifically, the focus is on the Space Shuttle Main Engine (SSME), the most powerful reusable liquid propellent rocket engine, built after many years of experience in designing expendable rockets. The SSME (Figure 1-1) provides a considerable part of the thrust for the Space Shuttle during lift-off and is driven to its limit during normal operation. Therefore, the reliable performance of the SSME is absolutely indispensable for successful missions. The large-scale SSME itself has a large number of parts working together in a complex way, but the only sources of information that can be extracted for monitoring its health and performance are a set of on-board sensors.

The failure detection problem for the SSME involves the detection and iso-
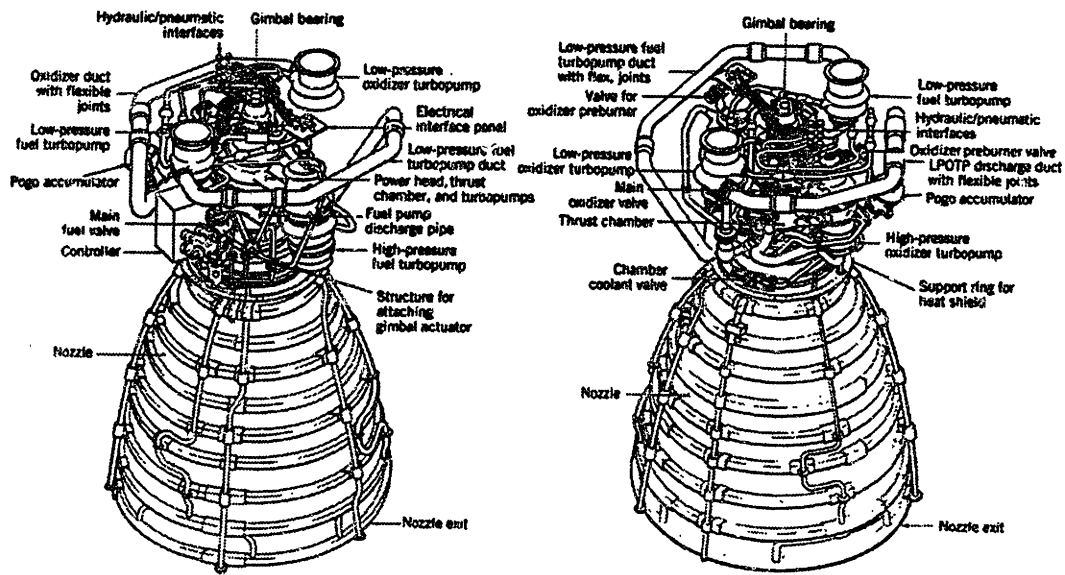
Figure 1-1: Two views of the Space Shuttle Main Engine (SSME)

lation of engine failures using the information measured by the sensors. The solution described in this thesis is a Vehicle Health Monitoring (VHM) system that detects and isolates failures in the engine control valves, sensors, and the internal components such as the turbopumps, the combustion chamber, and the main injector. This proposed VHM system uses model-based statistical signal processing methods. Specifically, using a dynamic model of the engine and its sensors, we design Kalman filters which predict the sensor outputs. The difference between the predicted and the observed measurements, called the residual, is then analyzed using detection and information theoretic methods.

1.1 Motivation

The present safety monitoring system [22] of reusable launch vehicle uses the so-called redlines system, where redlines are defined to be the limits at which the engine is inoperable. These limits correspond to the values at which selected measurements or derived core paramters exceed above their normal operating

14

values. While this approach is effective in avoiding catastrophic failures caused by a specific component, it has many drawbacks. For example, during the developmental and operational phases, the engine has already undergone severe damage by the time the failure is declared. Moreover, this technique causes high false alarm rate that usually results in premature shutdown of the engine due to normal excursions and inconsistent redlines limits that vary from one engine to another, since components of each engine are produced by different manufacturers.

In this thesis, we propose a Vehicle Health Monitoring(VHM) system which not only enhances the current redlines system, but also continuously supplies useful information for the operation of engine. Our VHM system provides information about the condition of the engine state variables that would help identify trends indicating incipient anomalies. This information allows the main controller to execute appropriate actions for system reconfigurations such as gain adjustment, or activation of backup system, etc. These actions reduce damage after a failure is detected. In addition, the information supplied by the VHM system can be used for reducing the cost of between-flight maintenance, and can aid in sensor placement and selection.

1.2 An Overview of Some Failure Detection Methods

Many approaches are potentially applicable for SSME health monitoring. In particular, for failure detection, some notable ones are redlines, pattern recognition and data trending, plume spectrometry, vibration monitoring, and model-based failure detection. There are of course overlaps of ideas in these techniques, and a comprehensive approach to FDI (Failure Detection and Isolation) for the SSME would have to integrate these and perhaps many other techniques into a framework that yields maximal consistency and reliability in detecting and isolating

failures. We will give a brief description of each technique.

Redlines is perhaps the oldest technique and the easiest to implement. Efforts have gone into designing more intricate redlines with improved criteria which incorporate tighter limits and account for variations in nominal values of the selected parameters due to different power levels and individual engine characteristics. These redline decision logics are usually part of an expert system. While these improvements seem to reduce the likelihood of extensive engine damage by the time the failure is declared, they are unreliable and may cause high false alarm rate if not implemented judiciously. Examples of usage of redlines are mentioned in [7, 8, 22, 23].

Pattern recognition and data trending are time series analysis methods that utilize the time history of the measurements of various failure scenarios to train a neural network or to define a specific subspace for each failure. The trained network performs FDI by identifying the signature of each particular failure. These techniques do not make use of the plant dynamics, and in general require large amount of data from the failed and nominal systems to sufficiently identify all failures. Considerable success in these techniques has been demonstrated in the works of [6, 10, 21].

Plume spectrometry can be independently implemented from other techniques with a set of simple sensors that have minimal interface with the engine. This technique examines the engine exhaust plume's combustion products since different combustion species possess different spectral bands. Therefore, analysis of the spectral emission, i.e. the amplitude and wavelength, can reveal the presence of anomalies and often lead to identify the failed component. A number of results have been developed for this technique in [5, 26].

Vibration monitoring of turbopumps is another important technique that provides early detection of the engine's structural failures. Analysis of the vibration data involves identifying and categorizing peaks in the power spectral densities that represent the fundamental and harmonic frequencies of the rotation in the shafts and the bearing cage. Anomalies are detected by observing

the amplitude of the peaks of the fundamental and harmonic frequencies. This technique has promising results as reported in [16].

The VHM system proposed in this thesis is a model-based failure detection technique which uses the RLV engine dynamics, together with sensor models, to design filters that monitor the physical state of the engine. These filters use sensor measurements together with a model of the engine dynamics to supply outputs that are affected in a specific way by each failure. A decision rule, via threshold test or likelihood ratio test, detects and isolates the failures based on these outputs. Approaches that exploit system dynamics include the work in [9, 14, 28]. An important difference in our work is the use of the entire dynamics of the SSME [1]. The dynamic model uses a set of 37 state variables and an assumed set of 15 sensors. In actual ground test setting, approximately 500 measurements are taken along with visual inspection of video cameras and technical staffs. In real-time operation, these ground-based facilities are not accessible and thus we limit our work to using a smaller set of sensors for obtaining information on the engine.

Furthermore, one significant strength of our model-based technique stems from the fact that the system dynamics provide additional physical insight into the interpretation of the failure signatures. For instance, when failures in two different sensors yield similar signatures, then knowledge of the physical relationship between these measured variables could provide a more definite conclusion, as we intend to demonstrate. In other words, the proposed method makes use of the tight interaction between the physics of the real engine, the engine's mathematical model, and the failure detection algorithm to check their mutual consistency.

1.3   VHM Design Concept

In this section, we present a VHM system that considers failures in the SSME sensors/valves and internal components. The VHM structure is shown in Figure
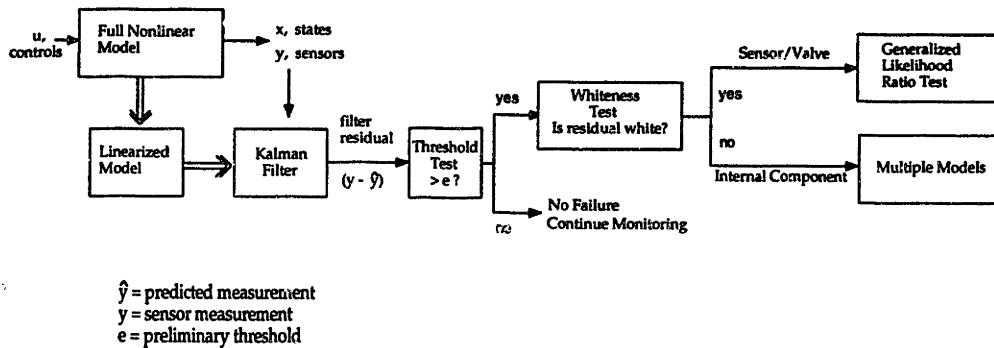
Figure 1-2: SSME Vehicle Health Monitoring System

1-2. First, nonlinear models of the engine available from Rocketdyne [1], are used to obtain linear models for any desired operating condition. The linearized models are then used to design Kalman filters which provide least-square estimates of the states in the engine model. The motivation of using the Kalman filter will be clear as we describe the function of other components of the VHM system.

The failure detection and isolation scheme takes as input the difference between the predicted outputs given by the Kalman filter and the observed measurements, referred to as the residual. In the absence of failure, the residuals are unbiased (of zero-mean), demonstrating agreement between the estimates and the observed measurements. In contrast, biased residuals are indicative of abnormal behavior or failures, and based on this fact, a *preliminary* threshold test is continuously used to detect bias. If the value of a residual exceeds this threshold, then a failure is declared.

The next step in the FDI scheme is to isolate the failed component. The isolation problem is solved in two levels. The first level categorizes the failure as either a sensor/valve failure or an internal component failure by checking the whiteness of the residual. As we will discuss later, a white residual indicates that the failure is more likely to have occurred in a sensor or a valve whereas a colored or correlated residual indicates that the failure is more likely to have occurred in an internal component, such as the turbopump or the main injector.

18

After the categorizing, the second level of isolation determines the specific sensor or valve, or internal component which has failed. For a sensor/valve failure, the Generalized Likelihood Ratio Test (GLRT) is applied to detect the presence of the failure with a more complex decision rule and decide which component has failed. On the other hand, if the residual is correlated, then the Multiple Model(MM) method is used in the decision process to identify the faulty subsystem. The GLRT and the MM method are two different kinds of hypothesis test as discussed in Chapters 3, 4, and 5.

Thus, the VHM system is mainly concerned with two functions: 1) generating the residual using Kalman filters and 2) making a decision based on this information. The threshold test determines whether a failure has occurred, while the whiteness test determines whether the failed part is among the valves and sensors, or among the internal components. Further isolation is accomplished using Generalized Likelihood Ratio Test for sensor/valve failures or the Multiple Model method for internal component failures. The logic behind our choice of the structure of the VHM and the failure detection sequence is justified in Chapter 3.

The motivation behind dividing the failures into two categories, one for sensors and valves and the other for internal components, is due to the fact that these failures are modeled in different fashions. Sensor and valve failures are modeled as additive changes at the output and input of the SSME, respectively. Additive changes are independent of the state and reflect our belief of how the real changes occur in the component in question. The usage of additive changes to represent failures in the sensors and valves has provided reasonable results, as it is a common way to represent bias in the gyros.

With the SSME, for instance, a defect in a sensor can be commonly represented as a bias, or increases in measurement noise. A valve fault can manifest as changes in the control matrix, or as a bias in the input due to a stuck valve. Internal component failures, on the other hand, are represented as multiplicative changes because they change the dynamic characteristic of the system. For

instance, a bearing failure in the engine's turbopump would shift the pump's dynamics to induce rubbing, and consequently affect the entire dynamics of the engine.

In the context of the FDI scheme, sensor/valve failures and internal component failures are distinguished by their effect on the correlation property of the residual. The isolation of these two types of failures is also at different levels of difficulty. Sensor or valve failures can be resolved by many failure detection techniques, such as the GLRT with high accuracy. Isolation of internal component failures, however, is a much more difficult problem, as the SSME is a highly coupled system, and there are few techniques available. Ideally the objective is to identify the exact origin of the failure. Unfortunately, determining which specific blade is faulty in a turbopump is very difficult and may require placing extra sensors in each pump, where the physical space is extremely limited. Isolation to the highest level of detail is rarely possible in any system, let alone a highly coupled one like the SSME. A realistic objective for the SSME is to determine in which of the following subsystems has a faulty component: the High Pressure Fuel Turbopump, the High Pressure Oxidizer Turbopump, the Low Pressure Oxidizer Turbopump, the Low Pressure Fuel Turbopump, the Main Injector, the nozzle, and the pipes/ducts.

Therefore, instead of attempting to detect and trace back the precise origin (component) of the failure, the VHM's algorithm specifies which subsystem exhibits anomalous behavior. For example, the algorithm identifies that the High Pressure Oxidizer Turbopump subsystem has failed, but not that a specific component (i.e. the turbine blade, or bearing) in that subsystem has malfunctioned. This approach is attractive and useful since internal component failures tend to quickly propagate from one component to another in the same subsystem before spreading to other subsystems. The knowledge of the degraded subsystem is crucial in guiding subsequent investigations such as using vibration monitoring to further analyze the identified faulty subsystem.

For the particular problem of designing a FDI scheme for the engine internal

component failures, using the MM technique is sufficient for the purpose of isolating failures on the level of subsystems. The context of subsystems refers to a collection of components in the engine. For instance, the High Pressure Oxidizer Turbopump Subsystem is composed of the pump itself, the turbine, the injector, and the oxidizer preburner. The division of the subsystems and the MM method are described in Chapter 5.

1.4   Thesis Outline

Chapter 2 presents the operational performance and dynamics of the SSME. Mathematical modeling of the dynamics of the engine are discussed in terms of the 37 nonlinear differential equations categorized by seven basic type of thermodynamics relations. The discrete-time linear state space representation of this model at various power levels as well as the sensor model are derived with noise included. The characteristics of the open loop dynamics of the engine are also discussed.

In Chapter 3, an overview and justification of the proposed VHM system illustrated in Figure 1-2 is presented. We focus on the Kalman filter design, the preliminary threshold test for detection, and the test of whiteness. This takes us to the first level of isolation, i.e. determining whether the failure is in a sensor or a valve, or in an internal subsystem.

In Chapter 4, we focus on the development of the residual generation and the failure detection process for sensor/valve failures. The Kalman filter described in Chapter 3 is used as the residual generator while the GLRT is used for failure detection and isolation for all 15 sensors and 5 control valves. Results are presented through numerical simulations. A geometrical interpretation of the GLRT is also included to provide additional insight into the detection problem.

In Chapter 5, we consider engine internal component failures. We discuss the different types of internal component failures and their characteristics, and how the proposed VHM algorithm performs FDI of these failures. The Multiple

Model method is derived and simulation results are presented for a few representative failures that occurred in the past. Finally the issue of robustness of the MM method is considered using information theoretic measures that give the *separation distance* among hypothesized models.

Chapter 6 provides conclusions and suggests further research to enhance the VHM algorithm presented in this thesis.

Chapter 2

The SSME: Description and Modeling

In this chapter, the operational performance specifications and mathematical modeling of the engine are outlined. More details about the SSME can be found in [12], [24], and [25]. We begin with providing the description of the engine in Section 2.1 and its mathematical modeling in Section 2.2. The linearization of the SSME nonlinear model for the analysis of failure detection is next discussed in Section 2.3, and the engine's dynamic characteristics are delineated in Section 2.4.

2.1   Description of the engine

The design and operation of the SSME represents a revolution in rocket propulsion technology. The power output of its four turbopumps drastically surpasses that of its predecessor; each SSME can produce about 12.3 million horsepower compared to that of 11000 horsepower by the J-2 engine flown on the Saturn V vehicle. The SSME operates at higher temperature and pressure extremes than any other mechanical system. The temperature and pressure in the main combustion chamber(MCC) reaches about 6000 degrees Farenheit and 3200 psi respectively at full power level. The highest MCC pressure achieved prior to
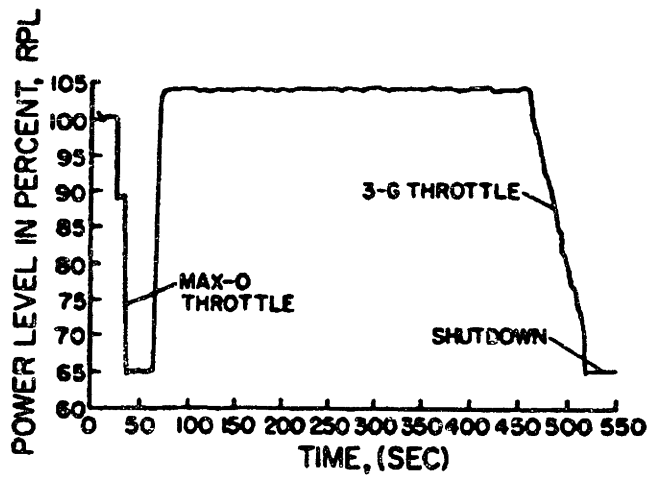
Figure 2-1: A typical SSME Thrust Profile
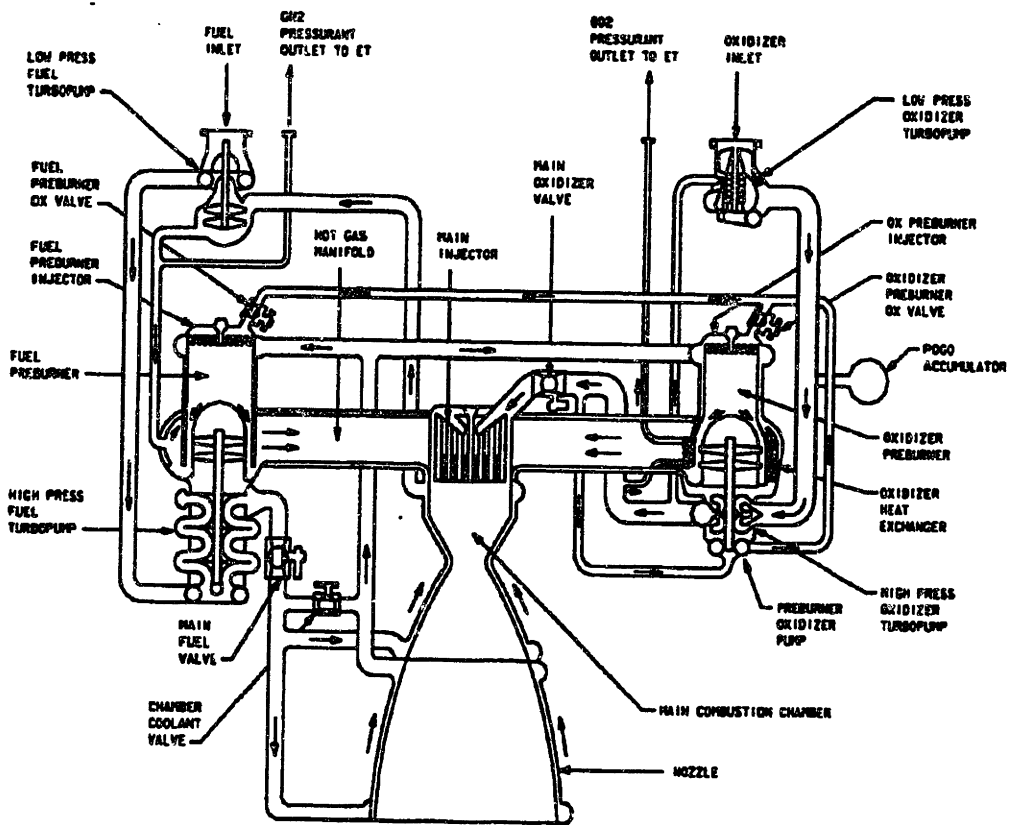
Figure 2-2: SSME Schematic Propellant Flow

SSME was only 1000 psi by the Saturn V engine. These achievements demonstrate a large jump in reusable space engine technology brought forth by the SSME.

The SSME uses cryogenic Liquid Hydrogen(LH) and liquid oxygen(LOX) as propellants. The engine power operation ranges from a minimum level of 305,000 lbs of thrust (65 percent) to Full Power Level (FPL), or 512,300 lbs of thrust (109 percent) in 1 percent increments. A typical operational profile of the SSME is illustrated in Figure 2-1. The major subsystems of the SSME consists of four turbopumps, one main injector (MI), one main combustion chamber (MCC), one nozzle, heat exchangers, valves, ducts. The four turbopumps are the Low-Pressure Oxidizer Turbopump (LPOTP), the Low-Pressure Fuel Turbopump (LPFTP), the High-Pressure Oxidizer Turbopump (HPOTP), and the High-Pressure Fuel Turbopump (HPFTP). Each of the two high pressure pumps has a preburner and an injector.

The SSME is a staged combustion cycle engine, which provides higher levels of performance when compared to other cycles, such as the bi-propellant gas generator cycle and expander cycle. The staged combustion cycle maximizes the power performance since all combustion products and propellents are eventually utilized to generate thrust. A schematic propellent flow for the engine is shown in Figure 2-2. The discharged LOX and LH from the pumps are first partially diverted to the preburners where they are combusted at a mixture ratio of one. This rich mixture ratio is required to keep the gaseous combusted product at moderate temperature to prevent from damaging the turbine blades while still satisfying the power required by pumps. The combusted products are next expanded through the turbines to drive the pumps, and then are directed to the main combustion chamber along with the coolant fuel and the rest of the LOX to go through a second combustion process at mixture ratio of 6 lb of LOX to 1 lb of LH.

The selection of 3000 psi chamber pressure and staged combustion cycle to attain maximum specific impulse for the SSME necessitates the use of four tur-

bopumps operating at high speed. The design selection of the four turbopumps are determined by the fact that low pressure is desirable at the pump inlet to minimize the tank's thickness, while high pressure is necessary at the pump outlet to produce discharge pressure between 7000 psi and 8500 psi. These conflicting requirements are best balanced by adding two low pressure pumps (one for LOX and one for LH) that give sufficient inlet pressure for the two high pressure pumps that, in turn, must provide high discharge pressure.

The SSME operational phases and control are accomplished by simultaneous manipulation of the five hydraulic actuated valves: Main Fuel Valve (MFV), Main Oxidizer Valve (MOV), Chamber Coolant Valve (CCV), Oxidizer Preburner Oxidizer Valve (OPOV), and Fuel Preburner Oxidizer Valve (FPOV).

In the start-up phase, all five valves are scheduled to ensure a smooth start by directing fuel into the combustion chamber prior to the oxidizer to establish a fuel-rich environment. This start-up sequence takes approximately 4 seconds.

In the main stage phase where the power level is about 104 percent as illustrated in Figure 2-1, the MOV, MFV, and CCV remain fully opened until shutdown. The OPOV and FPOV control engine throttling by varying the output of the preburners, the speed of the turbopumps, and consequently the LOX and LH flow rate. The mixture ratio in the MCC is solely controlled by the FPOV. The OPOV is used simultaneously with the FPOV to control the thrust at a specific mixture ratio.

2.2   Mathematical Modeling

This section presents the mathematical model of the dynamic behavior of the SSME. The dynamics of the SSME are governed by the interaction of a large number of physical variables. The SSME flow schematic of Figure 2-2 provides a notion of the potential degree of complexity of these interactions.

2.2.1 Nonlinear Model

The open loop dynamic model presented here is adapted from Rocketdyne's documentation and nonlinear simulation [1]. It will be referred to as Rocketdyne Model hereafter. The core of this model is a set of 37 nonlinear ordinary differential equation that can be written in a concise notation. Specifically,

$$\dot{z} = f(z(t), s(t)) \tag{2.1}$$

$$q(t) = h(z(t)) \tag{2.2}$$

where $z(t)$ is the 37 dimensional state vector, $s(t)$ is the vector representing the 5 inputs or valve openings, $q(t)$ is 15 dimensional observation vector representing sensor outputs at time t, and $h(z(t))$ is a nonlinear function of the observation. In this study, it is assumed that a total of 15 sensors are used to measure various state variables of the model in Eq.(2.1) and Eq.(2.2). While this is probably a subset of the entire actual set of sensors, it is sufficient to ensure the observability of every state in Eqs.(2.1-2.2). The description of the thirty-seven variables are listed in Table 2.1 with unit in paratheses. Ten of these 37 variables are sensed directly and 5 measurements are nonlinearly related to more than one variable, as listed in Table 2.2.

2.2.2 Seven Types of Equation

The 37 equations fall into seven categories of dynamic equations [24]:

- Type 1: *The rotational dynamics of turbopumps*: Four equations, one for each turbopump, describe the rotational dynamics of the engine. The following is the general form of this equation, where I_{TP} is the moment of inertia of the rotor, Ω the angular velocity, τ_T the torque delivered by the turbine, and τ_P the torque absorbed by the pump:

$$I_{TP}\frac{d\Omega}{dt} = \tau_T - \tau_P \tag{2.3}$$

27

State Variable	Description
Ω_{O2}	Rotational velocity (rad/sec) in HPOTP (sensor 1)
P_{POS}	Pressure (psia) for preburner common supply
$\dot{m}_{FPO}$	Mass flowrate (lb/sec) for oxidizer preburner oxidizer
$\dot{m}_{OPO}$	Mass flowrate (lb/sec) for oxidizer preburner oxidizer
P_{OP}	Pressure (psia) in oxidizer preburner
P_{FI}	Pressure (psia) for main fuel injector (sensor 3)
$\dot{m}_{MOV}$	Mass flowrate (lb/sec) for main oxidizer valve
P_C	Pressure (psia) in thrust chamber (sensor 4)
Ω_{F2}	Rotational velocity (rad/sec) in HPFTP (sensor 5)
$\dot{m}_{CLI}$	Mass flowrate (lb/sec) at coolant line inlet
$\dot{m}_{CLO}$	Mass flowrate (lb/sec) at coolant line outlet
P_{CL}	Pressure (psia) in coolant line
$\dot{m}_{FPF}$	Mass flowrate (lb/sec) into fuel preburner
P_{FP}	Pressure (psia) in preburner (sensor 6)
$\dot{m}_{OPF}$	Mass flowrate (lb/sec) into oxidizer preburner
$\dot{m}_{MC}$	Mass flowrate (lb/sec) in main chamber heat exchanger
T_{W15}	Temperature of hot gas wall at cooling line of the main chamber
T_{W25}	Temperature of ambient wall at cooling line of the main chamber
SU_4	Specific internal energy of hydrogen (Btu/lb) out of nozzle coolant line
$\dot{m}_{FN}$	Mass flowrate (lb/sec) at the primary heat exchanger
ρ_4	Mass density of hydrogen (lb/in^3) out of nozzle coolant line
T_{W14}	Temperature of hot gas wall at nozzle coolant line
T_{W24}	Temperature of ambient wall at nozzle coolant line
$\dot{m}_4$	Mass flowrate (lb/sec) through nozzle coolant line
$\dot{m}_{ASIFP}$	ASI Mass flowrate (lb/sec) in fuel preburner
$\dot{m}_{ASIOP}$	ASI Mass flowrate (lb/sec) in oxidizer preburner
$\dot{m}_{ASIMC}$	ASI Mass flowrate (lb/sec) in main chamber
P_5	Pressure (psia) in the preburner common fuel supply line
$\dot{m}_{FNBP}$	Mass flowrate (lb/sec) through primary nozzle by pass
$\dot{m}_{OT1}$	Mass flowrate (lb/sec) in low-pressure oxidizer turbine
P_{OS}	Pressure (psia) in low-pressure oxidizer pump inlet
Ω_{O1}	Rotational velocity (rad/sec) in LPOTP (sensor 9)
Ω_{F1}	Rotational velocity (rad/sec) in LPFTP (sensor 10)
$\dot{m}_{OS}$	Mass flowrate (lb/sec) through LPOTP inlet

Table 2.1: Complete SSME state variables (directly sensed variables indicated).

Sensor	Description
P_{od2}	HPOTP discharge pressure (psia) (sensor 11)
P_{od3}	Preburner booster pump discharge pressure (psia) (sensor 12)
T_{ft2d}	High pressure fuel turbine discharge temperature (R) (sensor 13)
T_{ot2d}	High pressure oxidizer turbine discharge temperature (R)(sensor 14)
P_{fd2}	High pressure fuel pump discharge pressure (psia) (sensor 15)

Table 2.2: Nonlinear sensors

- Type 2: *Inertia of liquid under pressure variation*: Seventeen equations follow the form of this type of equation. This equation relates the change in time in the mean flow rate at each pipe to the pressure difference. It is just an application of Newton's second law and is analogous to inductance effects in electric circuits. Consider a pipe of length L and cross-section area A, fed by a pump discharge at a pressure P_D, and having a mean pressure P_P . Frictional forces along the pipe at bends and restrictions contribute a total pressure drop of $\lambda \left(1/2\rho v^2\right)$, where ρ and v are the liquid density and velocity respectively, and λ (of order unity) is the pressure loss coefficient. From the continuity equation $\dot{m} = \rho v A$ where $\dot{m}$ is the mass flow rate, we can write

$$\left(\frac{L}{A}\right) \frac{d\dot{m}}{dt} = P_D - P_P - \lambda/2\rho A^2 \dot{m}^2 \qquad (2.4)$$

- Type 3:*Fluid capacitance under pressure variation*: This type of equation is analogous to capacitive effects. There are 8 equations of this type expressing the ability of cavities to store fluid due to its compressibility under pressure fluctuations. Consider the same pipe as above with an inlet flow rate of $\dot{m}_{in}$ from the pump, and discharges at one or more different places so that the net outlet flow is $\Sigma\dot{m}_{out}$. A net inflow is then given by $\Sigma\dot{m} = \dot{m}_{in} - \Sigma\dot{m}_{out}$. Although the fluid discharged by the pump is a liquid, it still has finite compressibility, ν, which becomes more appreciable as the

pressures become large. This is the case of the turbomachineries (both low pressure and high pressure) we are dealing with, and as a result the net inflow does not generally equal zero. From the definition of compressibility and considering that the volume V of the pipe remains constant under the pressure fluctuations, we can write the continuity equation as

$$\nu \rho V \frac{dP}{dt} = \Sigma \dot{m} \tag{2.5}$$

- Type 4:_Fluid Capacitance under density variation_: Some of the dynamics of the SSME, in 2 equations, are more sensitive to density variations than to pressure variations. This is specially true when we deal with heat transfer phenomena in the main combustion chamber and nozzle walls. In either case, this type of equation is essentially the same as Eq.(2.5) with constant pressure:

$$V \frac{d\rho}{dt} = \Sigma \dot{m} \tag{2.6}$$

- Type 5:_Heat transfer in the heat exchangers_: There are 2 equations that deal with the property of gases formed from heat transfer processes in the main heat exchangers of the engine. Consider the coolant fluid inside two walls separating the hot gases produced from the combustion process (from nozzle and combustion chamber) and the ambient gases of the atmosphere, and let u be its specific internal energy, h be its specific enthalpy and $\dot{m}h$ its heat flow rate convected by the system. From the first law of thermodynamics, assuming constant density, a simple heat flow balance for the coolant fluid yields

$$\rho V \frac{du}{dt} = \dot{Q}_{w,in} - \dot{Q}_{w,out} + (\dot{m}h)_{in} - (\dot{m}h)_{out} \tag{2.7}$$

where $\dot{Q}_{w,in}$ and $\dot{Q}_{w,out}$ are the heat flow rate transfered from the hot wall

to the fluid and the heat flow rate transfered from the fluid to the ambient wall respectively.

- Type 6: _Heat Transfer Across Metal Walls_: Four equations in the model describe the transient heat transfer phenomena inside the metal walls of the heat exchangers. Again, using the first law of thermodynamics, the net heat flux inside the wall is given by $\dot{Q}_{c,in}$ - $\dot{Q}_{w,in}$ where $\dot{Q}_{c,in}$ is the heat flux transfered from the hot gases to the hot wall and $\dot{Q}_{w,in}$ is the heat flux transfered from the wall to the coolant fluid. Then the temperature of the hot wall obeys the equation

$$\frac{dT_{hw}}{dt} = \frac{1}{m\bar{c}}(\dot{Q}_{c,in} - \dot{Q}_{w,in}) \tag{2.8}$$

where m is the mass per unit area of wall and $\bar{c}$ is the specific heat of the material.

- Type 7: _Time Delay Equation_: Some of the sensors used to monitor the behavior of the SSME do not have an immediate response to changes in the state of variables. This delay can be categorized as a first order time lag. For instance, if we have one of these sensors measuring mass flow rate and $\dot{m}_o$ is the real signal and $\dot{m}$ is the sensor output shifted by a small time constant ϵ, then the relationship between the two signals is

$$\frac{d\dot{m}(t)}{dt} = \frac{1}{\epsilon}(\dot{m}_o(t) - \dot{m}(t)) \tag{2.9}$$

2.2.3   The Nonlinear Differential Equations

With the geometry and other structural specifications of the engine specified, the complete 37 differential equations can be constructed.

$$\frac{1}{0.916}\frac{d\Omega_{o2}}{dt} = \tau_{ot2} - \tau_{op2} - \tau_{op3} \tag{2.10}$$

$$\frac{1}{100}\frac{d\dot{m}_{op3}}{dt} = P_{OD3} - \mathbf{P}_{pos} - 0.000813\dot{m}_{op3}^2 \tag{2.11}$$

$$\frac{1}{38120}\frac{d\mathbf{P}_{pos}}{dt} = \dot{m}_{op3} - \dot{m}_{fpo} - \dot{m}_{opo} - \dot{m}_{op2c} \tag{2.12}$$

$$\frac{1}{2}\frac{d\dot{m}_{fpo}}{dt} = \mathbf{P}_{pos} - \mathbf{P}_{fp} - 0.02488\left(\frac{\dot{m}_{fpo}}{[A/\bar{A}]_{fpv}}\right)^2 - 0.1948\dot{m}_{fpo}^2 \tag{2.13}$$

$$\frac{d\dot{m}_{opo}}{dt} = \mathbf{P}_{pos} - \mathbf{P}_{op} - 0.260\left(\frac{\dot{m}_{opo}}{[A/\bar{A}]_{opv}}\right)^2 - 1.463\dot{m}_{opo}^2 \tag{2.14}$$

$$\frac{1}{10000}\frac{d\mathbf{P}_{op}}{dt} = \dot{m}_{opf} + \dot{m}_{opo} - 0.9980\dot{m}_{ot2} \tag{2.15}$$

$$\frac{1}{3000}\frac{d\mathbf{P}_{fi}}{dt} = \dot{m}_{ft1} + \dot{m}_{ot2} + \dot{m}_{ft2} - 1.085\dot{m}_{fi} \tag{2.16}$$

$$\frac{1}{25}\frac{d\dot{m}_{mov}}{dt} = P_{od2} - \mathbf{P}_c - 0.001358\left(\frac{\dot{m}_{mov}}{[A/\bar{A}]_{mov}}\right)^2 - 0.0003573(\dot{m}_{mov})^2$$
$$\tag{2.17}$$

$$\frac{1}{4000}\frac{d\mathbf{P}_c}{dt} = \dot{m}_{fi} + \dot{m}_{mov} - \dot{m}_{cn} \tag{2.18}$$

$$\frac{1}{0.3087}\frac{d\Omega_{f2}}{dt} = \tau_{ft2} - \tau_{fp2} \tag{2.19}$$

$$\frac{1}{0.1}\frac{d\dot{m}_{cli}}{dt} = P_{fd2} - \mathbf{P}_{cl} - 6573.46\dot{m}_{cli}^2 \tag{2.20}$$

$$\frac{1}{0.1}\frac{d\dot{m}_{clo}}{dt} = \mathbf{P}_{cl} - \mathbf{P}_{fi} - 109.59\dot{m}_{clo}^2 \tag{2.21}$$

$$\frac{1}{283700}\frac{d\mathbf{P}_{cl}}{dt} = \dot{m}_{cli} - \dot{m}_{clo} \tag{2.22}$$

$$\frac{1}{20}\frac{d\dot{m}_{fpf}}{dt} = \mathbf{P}_9 - \mathbf{P}_{fp} - 0.000131\frac{\dot{m}_{fpf}^2}{\rho_9} \tag{2.23}$$

$$\frac{1}{16000}\frac{d\mathbf{P}_{fp}}{dt} = \dot{m}_{fpf} + \dot{m}_{fpo} - 0.9908\dot{m}_{ft2} \tag{2.24}$$

$$\frac{1}{10}\frac{d\dot{m}_{opf}}{dt} = \mathbf{P}_9 - \mathbf{P}_{op} - 0.000570\frac{\dot{m}_{opf}^2}{\rho_9} \tag{2.25}$$

$$\frac{1}{0.001}\frac{d(\rho_5\mathbf{SU5})}{dt} = \dot{Q}_{15} + \dot{Q}_{25} + 189.7\dot{m}_{mc} - \dot{m}_5 H_5 \tag{2.26}$$

$$\frac{1}{5}\frac{d\dot{m}_{mc}}{dt} = P_{mfvd} - P_5 - 0.001442\frac{\dot{m}_{mc}^2}{\rho_5} \tag{2.27}$$

$$\frac{1}{0.001}\frac{d\rho_5}{dt} = \dot{m}_{mc} - \dot{m}_5 \tag{2.28}$$

$$\frac{1}{0.31}\frac{d\mathbf{T}_{w15}}{dt} = \dot{Q}_{tc5} - \dot{Q}_{15} \tag{2.29}$$

$$\frac{1}{0.0825}\frac{d\mathbf{T}_{w25}}{dt} = -\dot{Q}_{25} \tag{2.30}$$

$$\frac{1}{0.0005}\frac{d(\rho_4\mathbf{SU4})}{dt} = \dot{Q}_{14} + \dot{Q}_{24} + 166.8\dot{m}_{fn} - \dot{m}_4 H_4 \tag{2.31}$$

$$\frac{1}{15}\frac{d\dot{m}_{fn}}{dt} = P_{mfvd} - P_4 - C_{\dot{m}_{fn}}\dot{m}_{fn}^2 \tag{2.32}$$

$$\frac{1}{0.0005}\frac{d\rho_4}{dt} = \dot{m}_{fn} - \dot{m}_4 \tag{2.33}$$

$$\frac{1}{0.1668}\frac{d\mathbf{T}_{w14}}{dt} = \dot{Q}_{tc4} - \dot{Q}_{14} \tag{2.34}$$

$$\frac{1}{0.0834}\frac{d\mathbf{T}_{w24}}{dt} = -\dot{Q}_{24} \tag{2.35}$$

$$\frac{1}{21.19}\frac{d\dot{m}_4}{dt} = P_4 - P_9 - C_{\dot{m}_4}\frac{\dot{m}_4^2}{\rho_4} \tag{2.36}$$

$$\frac{1}{0.1}\frac{d\dot{m}_{asifp}}{dt} = P_{mfvd} - \mathbf{P}_{fp} - 1072.88\dot{m}_{asifp}^2 \tag{2.37}$$

$$\frac{1}{0.1}\frac{d\dot{m}_{asiop}}{dt} = P_{mfvd} - \mathbf{P}_{op} - 1025.04\dot{m}_{asiop}^2 \tag{2.38}$$

$$\frac{1}{0.1}\frac{d\dot{m}_{asimc}}{dt} = P_{mfvd} - \mathbf{P}_c - 3552.88\dot{m}_{asimc}^2 \tag{2.39}$$

$$\frac{1}{10000}\frac{d\mathbf{P}_9}{dt} = \dot{m}_4 + \dot{m}_{fnbp} - \dot{m}_{opf} - \dot{m}_{fpf} \tag{2.40}$$

$$\frac{1}{0.2}\frac{d\dot{m}_{fnbp}}{dt} = P_{mfvd} - \mathbf{P}_9 - 0.1297\left(\frac{\dot{m}_{fnbp}}{[A/\bar{A}]_{ccv}}\right)^2 \tag{2.41}$$

$$\frac{1}{1.2}\frac{d\dot{m}_{ot1}}{dt} = P_{od2} - P_{od1} - (.1142 + R_{ot1})\dot{m}_{ot1}^2 \tag{2.42}$$

$$\frac{1}{150}\frac{d\mathbf{P}_{os}}{dt} = \dot{m}_{os} - \dot{m}_{mov} - \dot{m}_{op3} \tag{2.43}$$

$$\frac{1}{0.37566}\frac{d\Omega_{o1}}{dt} = \tau_{ot1} - \tau_{op1} \tag{2.44}$$

$$\frac{1}{1.0449}\frac{d\Omega_{f1}}{dt} = \tau_{ft1} - \tau_{fp1} \tag{2.45}$$

$$\frac{1}{21.57}\frac{d\dot{m}_{os}}{dt} = P_{ot} - \mathbf{P}_{os} - .0112\dot{m}_{os} \tag{2.46}$$

In the above equations, the state variables are shown in **bold**. The other variables on the right hand side are algebraically related to either the state variables or control variables by different functions listed in Appendix B.

2.3 Linear State Space Representation for the SSME Model

The nonlinear dynamics of the 37 differential equations are made more tractable for the analysis of failure detection by linearizing them about various operating points. The linearized model is extremely useful for constructing linear filters as will be seen in Chapter 3. The linearization procedure is described in Appendix C. The plant and sensor models of Eqs.(2.1-2.2), when linearized and discretized in the absence of a failure about certain power level or constant operating point z_o and s_o, become

$$x(k+1) = Ax(k) + Bu(k) + \Gamma w(k) \qquad (2.47)$$

$$y(k) = Cx(k) + v(k) \qquad (2.48)$$

where k is time, $x(k)$, $u(k)$, and $y(k)$ are the discrete version of $z(t) - z_o$, $s(t) - s_o$, and $q(t) - q_o$ respectively. The matrices A, B, C are the Jacobians of Eqs.(2.1-2.2) evaluated at z_o and s_o. The constant operating point (z_o and s_o) results in these matrices being time-invariant as described in Appendix C. The vector $w(k)$ represents the process noise independent of $x(k)$, and $v(k)$ represents the sensor or measurement noise. It is also assumed that the $w(k)$ and $v(k)$ are independent, and $w(k)$ and $v(k)$ are zero-mean and Gaussian white noise sequences with

$$E(w(k)w'(k)) = Q(k) \qquad (2.49)$$

$$E(v(k)v'(k)) = R(k) \qquad (2.50)$$

The terms $\Gamma w(k)$ represent the exogenous disturbance as well as the model uncertainty, such as that resulting from the linearization. The matrix Γ is therefore a design choice. For simplicity, we also make the noise covariance matrices constant, i.e. $Q(k) = Q$ and $R(k) = R$. Numerical values of these

matrices and other relevant matrices at 100 percent power level are listed in Appendix A. We also use these matrices as the representative system matrices to generate the results in the subsequent chapters, similar results can be obtained for system matrices at other power levels. In addition, the sensor sampling rate is assumed to be 1000 Hz.

The above equations are used to design the Kalman filter and the FDI algorithm that are described in the next chapter.

2.4   Characteristics of the SSME Model

In this section, a few important characteristics of the SSME dynamics is delineated to further enhance our insights about the choice of the detection scheme described in the next chapters.

The trajectories of the state variables of the nonlinear model described in the previous sections at 100 percent Rated Power Level (RPL) are shown in Figure 2-3 where the x-axis is time and the y-axis is the value of each of the 37 state variables. The nonlinear model is initialized at a point in the neighborhood of the 100 percent power level steady-state value. The plots in Figure 2-3 indicate that the SSME is a highly damped system, since most of the variables reach steady state in less than 1 second. The oscillatory behavior in the states P_{os} (pressure in low-pressure oxidizer turbopump inlet) and dM_{os} (mass flowrate through LPOTP) is primarily due the compressibility of LOX, which is transferred from the LOX tank to the LPOTP in a pipe whose length is much larger than the other pipes' length. This long pipe sustains low frequency oscillation and effectively possesses capacitive behavior that P_{os} and dM_{os} exhibit.

The stability of the dynamics is also reflected from the eigenvalues of the system transition matrix A across the 100 percent power level where all of the eigenvalues are located inside the unit circle. A plot of the eigenvalues is shown in Figure 2-4. The rapid transient response of the state variables motivates the usage of the stationary Kalman filter for failure detection and isolation.
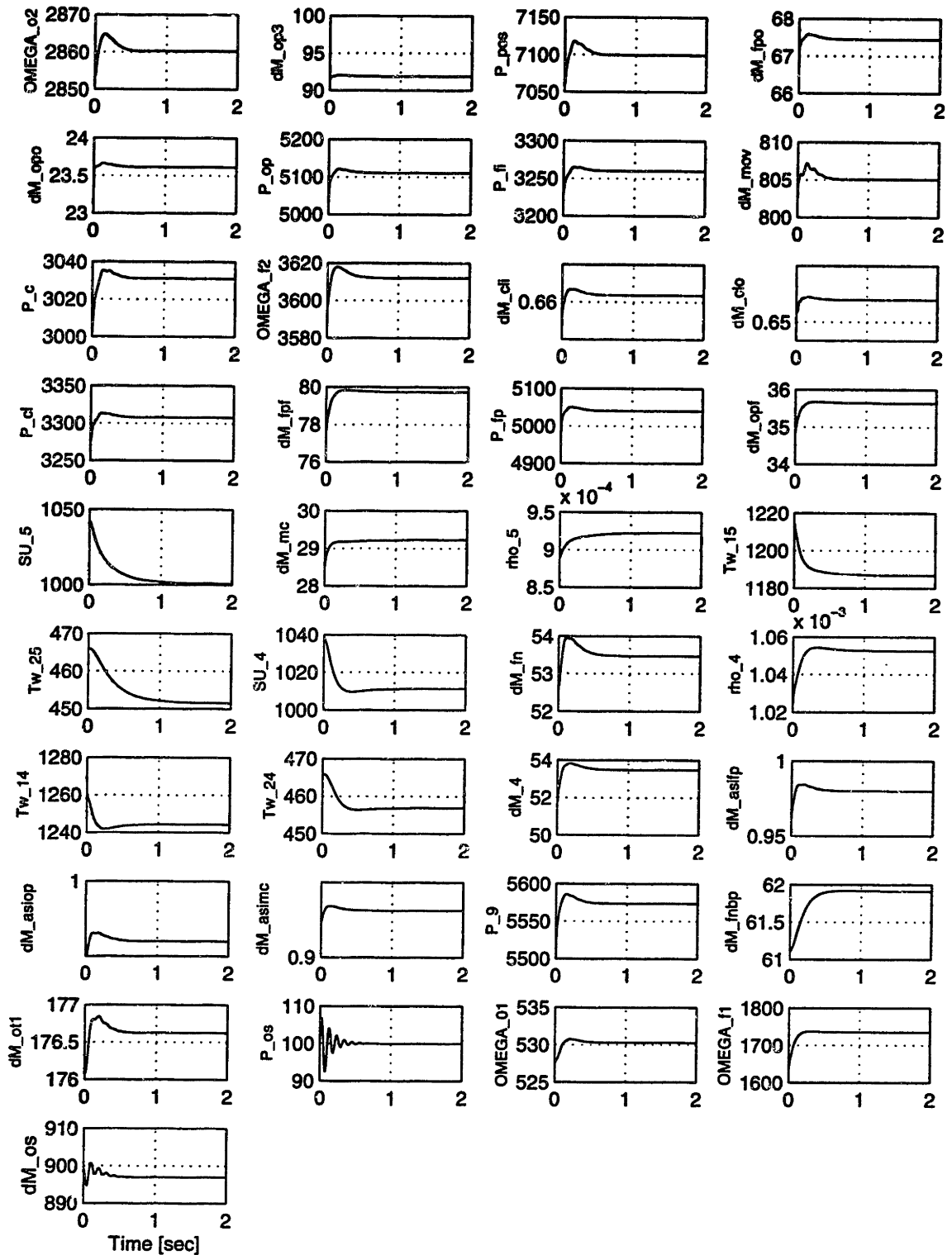
Figure 2-3: Trajectory of SSME model state variables when reaching steady-state at the 100 percent power level.
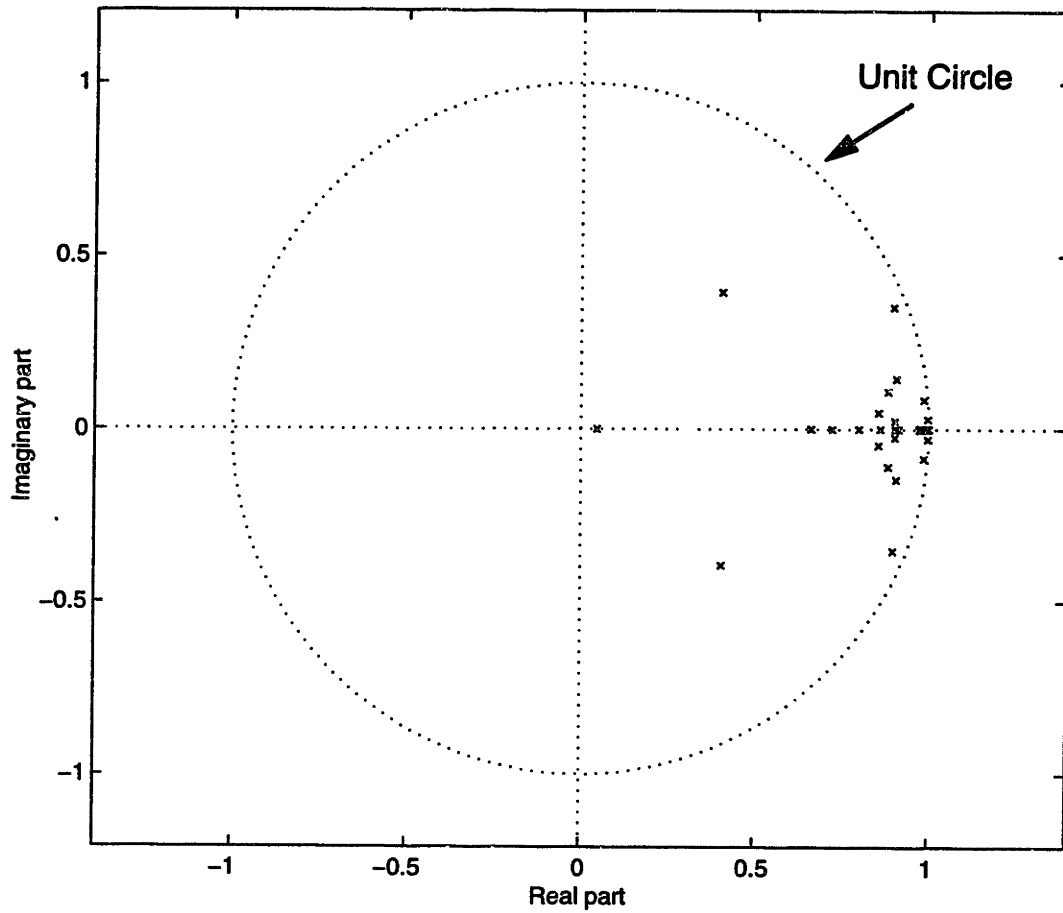
Figure 2-4: Pole-Zero Plot of Eigenvalues of A (in discrete time)

Chapter 3

The FDI Approach

This chapter motivates and delineates the role of each building block in the VHM system of Figure 3-1. In addition, a detailed description of the three of these blocks (in the dashed-line box) are presented: the Kalman filter, the threshold test, and the whiteness test.

In Section 3.1, the FDI problem is formulated using the nonlinear model. The difficulty of this formulation is explained, thus motivating the architecture in Figure 3-1. In Section 3.2, we reformulate the problem using the linear model introduced in Chapter 2 and explain the role of each of the subsequent blocks in the VHM system. Next, in Section 3.3, the linearized model is used to construct the Kalman filter (KF), and the vital role of the KF as a residual generator in

$\hat{y}$ = predicted measurement
y = sensor measurement
e = preliminary threshold
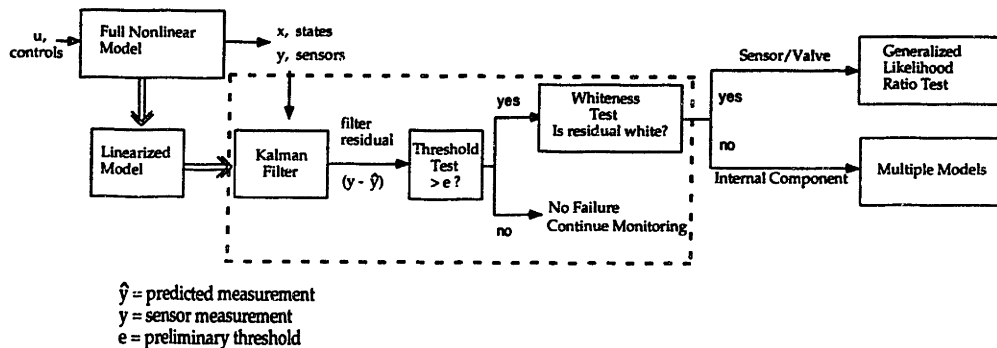
Figure 3-1: SSME Vehicle Health Monitoring System

the FDI process is demonstrated. In Section 3.4, we discuss the threshold test for failure detection. The whiteness test is introduced next in Section 3.5. This test is used to determine whether the detected failure occurred in a sensor or a valve, or in an internal component.

3.1   Problem Formulation and Solution

A natural approach to the problem of detecting and isolating failures in the SSME's sensors, valves, and internal components based on the observed measurements is to set up the problem as a hypothesis test. In this approach, the test consists of two hypotheses H_0 and H_1, where H_0 corresponds to the no-failure scenario and H_1 corresponds to the failure scenario. We assume that the failure either exists or does not exist in the entire time interval of interest, hence the time at which the failure occurs is not considered. The two hypotheses differ in the system parameter values used in the different hypothesized models. Let Θ_0 be the system parameter vector in the absence of any failure, and Θ_1 be the system parameter vector in the presence of a failure. Then, using the nonlinear model (2.1-2.2) with noise incorporated, the two hypotheses (in discrete time) are:

$$
\begin{aligned}
H_0 \ : \quad z(k+1) \ &= \ f(z(k), s(k); \Theta_0) + w(k) \\
q(k) \ &= \ h(z(k); \Theta_0) + v(k)
\end{aligned}
\tag{3.1}
$$

$$
\begin{aligned}
H_1 \ : \quad z(k+1) \ &= \ f(z(k), s(k); \Theta_1) + w(k) \\
q(k) \ &= \ h(z(k); \Theta_1) + v(k)
\end{aligned}
\tag{3.2}
$$

where $w(k)$ and $v(k)$ are zero-mean white Gaussian noise sequences, independent of each other with respective covariances Q and R. It is also assumed that $w(k)$ is independent with the state $z(k)$.

As shown in [27, 30], the usual procedure used to decide between H_0 and

H_1 is to compare a likelihood ratio (LR) $\bar{\Lambda}_{k_f}$ against a threshold η. Specifically,

$$\bar{\Lambda}_{k_f} = \frac{p(q(k_0), ..., q(k_f) \mid H_1, \Theta_1)}{p(q(k_0), ..., q(k_f) \mid H_0, \Theta_0)} \gtrless \eta \qquad (3.3)$$

Since the parameter vector Θ_1 is unknown, we use instead its maximum likelihood estimate $\hat{\Theta}_1$ assuming that hypothesis H_1 is true.

$$\hat{\Theta}_1 = \arg\max_{\Theta_1} p(q(k_0), ..., q(k_f) | H_1, \Theta_1) \qquad (3.4)$$

The estimate $\hat{\Theta}_1$ is then substituted into Eq.(3.3) to form the generalized likelihood ratio (GLR), $\Lambda_{k_f}(\hat{\Theta}_1)$. For an observation interval of $k_0 \leq k \leq k_f$, $\Lambda_{k_f}(\hat{\Theta}_1)$ is given by

$$\Lambda_{k_f}(\hat{\Theta}_1) = \frac{p(q(k_0), ..., q(k_f) \mid H_1, \hat{\Theta}_1)}{p(q(k_0), ..., q(k_f) \mid H_0, \Theta_0)} \gtrless \eta \qquad (3.5)$$

This overall procedure is referred to as the generalized likelihood ratio test (GLRT).

However, the GLRT as presented is impossible to compute for several reasons. First, the probability density function $p(q(k_0), ..., q(k_f))$ is usually not computable because of the nonlinearity in the model (2.1-2.2). Second, even if the density function can be derived, the sequence $\{q(k_0), ..., q(k_f)\}$ is not independent so that expression (3.5) cannot be computed recursively. Third, since the density function in Eq.(3.4) is generally not convex, the optimization in Eq.(3.4) is not simple.

To implement the GLRT, we use the approach illustrated in Figure 3-1. The outline of this approach is the following six steps:

1. Linearize the nonlinear model (Linearized Model Block).

2. Design a Kalman filter based on the linearized model in the absence of failure to generate a residual sequence (Kalman filter Block) that can be assumed independent.

3. Use the threshold test to process the residual and detect the presence of

40

failures (Threshold Test Block).

4. Perform the whiteness test to distinguish between sensor/valve failures and internal component failures (Whiteness test Block).

5. If the failure is in a sensor/valve, a simplified version of the GLRT is employed to detect and isolate the faulty component (GLRT Block).

6. If the failure is in an internal component, the Multiple Model (MM) method is employed to detect and isolate the faulty subsystem (MM Block).

In **Step 1**, the linearized model allows us to replace the density function $p(q(k_0), ..., q(k_f))$ with the density function $p(y(k_0), ..., y(k_f))$. The sequence $\{y(k_0), ..., y(k_f)\}$ is Gaussian since, by assumption, it is the output of a linear system driven by white Gaussian noise. The calculation of $p(y(k_0), ..., y(k_f))$, however, is still computationally intensive because the sequence $\{y(k_0), ..., y(k_f)\}$ is not an independent sequence which makes it impossible to express the GLRT recursively.

In the case when the failure occurs in a sensor or valve, a recursive form of the GLR can be obtained by replacing the sequence $\{y(k_0), ..., y(k_f)\}$ with its residual sequence $\{\varrho(k_0), ..., \varrho(k_f)\}$, which contains the same information as $\{y(k_0), ..., y(k_f)\}$ but has independent samples. The residual $\varrho(k)$ is defined as the difference between the output measurement $y(k)$ and the estimate of $y(k)$, $\hat{y}(k)$. The residual sequence is generated by a Kalman filter in **Step 2** based on the linearized model in the absence of failure.

Step 3 is concerned with failure detection. It consists of a test in which a statistic that is a function of the residual is continuously compared to a threshold to determine whether any a failure has occured. The threshold test eliminates other unnecessary computations (i.e. the whiteness, the GLRT, the MM method) if a failure is not detected. Details of the threshold test is discussed in Section 3.4.

41

Once a failure is detected, we proceed to **Step 4** where the whiteness test is used to determine whether the failure occured in a sensor or a valve, or in an internal component. This test takes advantage of the fact that the residual is white or uncorrelated when the failure is in a sensor/valve and is colored or correlated when the failure is in an internal component. The discussion of the whiteness test is in Section 3.5.

After the failure is detected and diagnosed either as a sensor/valve or internal component failure, then the algorithm proceeds as in Figure 3-1. Specifically, if it is determined that the failure is in a sensor or a valve, we proceed to **Step 5** and use the GLRT to specify which sensor or valve failed. If, however, the whiteness test indicates that the failure occured in an internal component, the MM method of **Step 6** is employed to identify the faulty subsystem. The GLRT is introduced in Section 3.2.1 and described in details in Chapter 4 while Section 3.2.2 introduces the MM method and Chapter 5 presents it in depth.

Before we proceed further, we want to note that the assumption of white Gaussian noise is an idealization since it allows us to obtain closed form solution for the GLRT. The white Gaussian noise is only an approximation of the observed realistic noises, hence this assumption should be kept in mind when implementing this algorithm on the real system.

3.2   Hypothesis Test with Linearized Model

In this section, we revisit problems with the GLRT presented in the last section. Specifically, we replace the nonlinear model with the linearized model and elaborate on the methods used in the treatment of both sensor/valve failures in **Step 5** and internal component failures in **Step 6**.

As a result of using the linearized model, the hypotheses of Eqs.(3.1-3.2) become

$$H_0 \quad : \quad x(k+1) \;=\; A(\Theta_0)x(k) + B(\Theta_0)u(k) + \Gamma w(k)$$

$$y(k) = C(\Theta_0)x(k) + v(k) \qquad (3.6)$$

$$H_1 : \quad x(k+1) = A(\Theta_1)x(k) + B(\Theta_1)u(k) + \Gamma w(k) + F(\Theta_1)\nu$$

$$y(k) = C(\Theta_1)x(k) + v(k) + L(\Theta_1)\nu \qquad (3.7)$$

for $k\epsilon[k_o, k_f]$. The parameter vector Θ_1 is the system parameters in the presence of internal component failures. The vectors F and L describe the way the failures in the control valves and sensors are injected into the system. For instance, if the failed component is a control valve, then F is the column vector of B representing that control valve in the plant dynamics. The magnitude of the sensor/valve failures is represented by ν, and ν is zero in the case of an internal component failure. Similarly, Θ_1 equals Θ_0 in the case of an sensor/valve failure. We want to point out here that the modeling of sensor and valve failures as additive changes in Eq.(3.7) implies that the magnitude of the failure, ν, is not a function of the state x. This modeling of failures as additive changes is generally valid since sensor and valve failures tend to occur independently of the behavior of the dynamics. While we only consider step failure mode for simplicity, other non-step failure modes can be analyzed in the same manner.

3.2.1   Sensor or Valve Failures

Let us first consider the case when the failure occurs in a sensor or a valve, i.e. $\Theta_1 = \Theta_0$ and $\nu \neq 0$. The hypotheses in Eqs.(3.6-3.7) become

$$H_0 : \quad x(k+1) = Ax(k) + Bu(k) + \Gamma w(k)$$

$$y(k) = Cx(k) + v(k) \qquad (3.8)$$

$$H_1 : \quad x(k+1) = Ax(k) + Bu(k) + \Gamma w(k) + F\nu$$

$$y(k) = Cx(k) + v(k) + L\nu \qquad (3.9)$$

where the matrices A, B, C are function of Θ_0. For an observations over a finite interval $k_0 \leq k \leq k_f$, the generalized likelihood ratio, $\Lambda_{k_f}(\hat{\nu})$, of the hypothesis

test of Eqs.(3.8-3.9) is given by

$$\Lambda_{k_f}(\hat{\nu}) = \frac{p(y(k_0), ..., y(k_f) \mid H_1, \hat{\nu})}{p(y(k_0), ..., y(k_f) \mid H_0)} \qquad (3.10)$$

$$\hat{\nu} = \arg \max_{\nu} p(y(k_0), ..., y(k_f) \mid H_1, \nu) \qquad (3.11)$$

where $\hat{\nu}$ is the maximum likelihood estimate of ν.

The likelihood ratio in Eq.(3.10) is computable since the output $y(k)$ is now the output of a linear system driven by white Gaussian noise; specifically the noise sequences $w(k)$ and $v(k)$. The computation of the LR itself, however, is not trivial. The difficulty primarily arises in the calculation of the density function $p(y(k_0), ..., y(k_f))$ because the sequence $\{y(k_0), ..., y(k_f)\}$ is not independent. As a result, the GLR cannot be written recursively. To overcome this difficulty, we implement a Kalman filter and replaces the measurement sequence $\{y(k_0), ..., y(k_f)\}$ by its residual sequence $\{\varrho(k_0), ..., \varrho(k_f)\}$ that carries the same information as the measurement. The $\varrho(k)$'s are statistically independent, allowing Eq.(3.10) to be computed recursively as we will show in depth in Chapter 4 where the we continue discussing sensor and valve failures in details.

3.2.2   Internal Component Failures

For the case when the failure is in an internal component, we have $\nu = 0$ and $\Theta_1 \neq \Theta_0$. The GLR and the optimization over Θ_1 is given by

$$\Lambda_{k_f}(\hat{\Theta}_1) = \frac{p(y(k_0), ..., y(k_f) \mid H_1, \hat{\Theta}_1)}{p(y(k_0), ..., y(k_f) \mid H_0)} \qquad (3.12)$$

$$\hat{\Theta}_1 = \arg \max_{\Theta_1} p(y(k_0), ..., y(k_f) \mid H_1, \Theta_1) \qquad (3.13)$$

The density function in Eq.(3.13) remains Gaussian but the mean and covariance are unknown because we don't know the system matrices $A(\Theta_1)$, $B(\Theta_1)$, and $C(\Theta_1)$ that generate the measurement $y(k)$. Moreover, even if the the density function can be derived, Eq.(3.12) cannot be expressed recursively since the

44

residual sequence is not independent. Thus the GLR in Eq.(3.12) is not computable. To circumvent these problems, we implement a bank of filters based on different hypotheses H_j, $j = 1, ..., N$

$$
\begin{aligned}
H_j \; : \quad x(k+1) &= A(\Theta_j)x(k) + B(\Theta_j)u(k) + \Gamma w(k) \\
y(k) &= C(\Theta_j)x(k) + v(k)
\end{aligned} \tag{3.14}
$$

where each filter assumes a specific model that corresponds to a particular internal component failure. These filters take as input the the measurement from the true failed model; and using the residual of each filter, the probability of hypothesis/model j is the correct hypothesis/model can be computed using Bayes rule. This is the MM method which we will elaborate in Chapter 5.

In the outline, before the GLRT or the MM method is used to isolate the failures in **Step 5** and **Step 6**, the failure must be detected by the threshold test in **Step 3** and categorized as either a sensor/valve failure or an internal component failure by the whiteness test in **Step 4**. Since both the threshold test and the whiteness test require the information from the residual or innovation, we will first describe how the residual is generated by the Kalman filter in **Step 2** and then proceed to **Step 3** and **Step 4**.

3.3   The Kalman Filter as a Residual Generator

In our FDI framework, the residual generated by the Kalman filter is an essential piece of information. In our VHM system, the residual is used in every phase of the FDI algorithm: from detecting the failure by the threshold test, to categorizing the type of failure (sensor/valve or internal component) by the whiteness test, and to isolating the failures by either the GLRT or the MM method.

In this section we present how the residual is generated by the Kalman filter algorithm. In addition, we discuss some of the powerful features of the residual,

45

namely its zero-mean and white (independent) properties. The derivation of the Kalman filter is given in Appendix C.

Consider the dynamic system (of Eqs.(2.47 -2.48)), rewritten below

$$x(k+1) = Ax(k) + Bu(k) + \Gamma w(k) \tag{3.15}$$
$$y(k) = Cx(k) + v(k) \tag{3.16}$$

where $w(k)$ represents the observation noise independent of $x(k)$, Γ is a known matrix , and $v(k)$ represents the measurement noise. It is also assumed that the $w(k)$ and $v(k)$ are independent, and $w(k)$ and $v(k)$ are zero mean, Gaussian white noise sequences with

$$E(w(k)w'(k)) = Q \tag{3.17}$$
$$E(v(k)v'(k)) = R \tag{3.18}$$

In addition, the initial condition $x(k_0)$ has mean $m_x(k_0)$ and covariance $P_x(k_0)$.

The Kalman filter gives the minimum variance estimates of $x(k)$ or the conditional means:

$$\hat{x}^-(k) = E[x(k)|y(k_0), ..., y(k-1)] \tag{3.19}$$
$$\hat{x}^+(k) = E[x(k)|y(k_0), ..., y(k)] \tag{3.20}$$

The estimates $\hat{x}^-(k)$ and $\hat{x}^+(k)$ are referred to as the predicted and update estimates of the state, respectively. The equations for the KF algorithm is as follows:

Prediction

$$\hat{x}^-(k+1) = A\hat{x}^+(k) + Bu(k) \tag{3.21}$$
$$P^-(k+1) = AP^+(k)A'(k) + \Gamma Q \Gamma' \tag{3.22}$$

Update

$$\hat{x}^+(k+1) = \hat{x}^-(k+1) + K(k)\varrho(k+1) \qquad (3.23)$$

$$\varrho(k) = y(k) - C\hat{x}^-(k) \qquad (3.24)$$

$$K(k) = P^-(k)C'(k)V(k)^{-1} \qquad (3.25)$$

$$V(k) = CP^-(k)C'(k) + R \qquad (3.26)$$

$$P^+(k) = P^-(k-1) - K(k)CP^-(k-1) \qquad (3.27)$$

where $P^-(k)$ and $P^+(k)$ are error covariance matrices of $\hat{x}^-(k)$ and $\hat{x}^+(k)$ respectively . The vector $\varrho(k)$ is referred as to the innovation in the absence of failure, and it is referred as the the residual in the presence of any failure. The matrix $K(k)$ is known as the Kalman gain and $V(k)$ is the covariance of the residual. If the measurement is available at time k_0, the algorithm is initiated with $\hat{x}^-(k_0) = m_x(k_0)$ and $P^-(k_0) = P_x(k_0)$, otherwise initiated with $\hat{x}^+(k_0) = m_x(k_0)$ and $P^+(k_0) = P_x(k_0)$.

A number of observations can be made regarding the structure of the Kalman filter. First, Eq.(3.21) is a noise-free version of Eq.(3.15) and its correction based on the previous measurement is given by Eq.(3.23). Second, the covariances (3.22 and 3.25-3.27) can be computed off-line since the measurement does not enter into these equations. Finally, in order to process the next estimate, only the current measurement instead of the whole past history of the measurement is needed for the computation.

In the context of FDI, the most desirable feature of the innovation in the absence of failure is that it is both white and of zero mean. Specifically,

$$E[\varrho(k)] = E\{y(k) - C\hat{x}^-(k)\} \qquad (3.28)$$

$$= E\{y(k) - E[y(k)|y(k_0),...,y(k-1)]\}$$

$$= E[y(k)] - E(E[y(k)|y(k_0),...,y(k-1)])$$

$$= E[y(k)] - E[y(k)] = 0 \qquad (3.29)$$

47

where the law of iterated expectation is applied to go from the third equality to the fourth equality. For $n < k$,

$$
\begin{aligned}
E[\varrho(n)\varrho'(k)] &= E\{E[\varrho(n)\varrho'(k)|y(k_0),...,y(n)]\} \\
&= E\{\varrho(n)E[\varrho(k)|y(k_0),...,y(n)]\} \\
&= 0
\end{aligned}
\tag{3.30}
$$

where the second equaltity follows from the fact that, given measurement through n, then $\varrho(n)$ is a constant. Similar arguments can be made for $k > n$. The third equality follows from

$$
\begin{aligned}
E[\varrho(k)|y(k_0),...,y(n)] &= E[y(k)|y(k_0),...,y(n)] - E\{E[y(k)|y(k_0),...,y(n)]\} \\
&= E[y(k)|y(k_0),...,y(n)] - E[\hat{y}(k)|y(k_0),...,y(n)] \\
&= E[y(k)|y(k_0),...,y(n)] - E[y(k)|y(k_0),...,y(n)] \\
&= 0
\end{aligned}
\tag{3.31}
$$

Thus the innovations (or residuals) are uncorrelated or white, which in the Guassian context, implies that the innovations are *independent*. This property of the innovation is used extensively in our application of determining whether a failure occured in a sensor/valve or in an internal component as will be discussed in Section 3.5. It is also used in Chapter 4 and 5.

3.4 The Threshold Test

As mentioned earlier, the residual (or innovation) generated by the KF is next processed by a preliminary threshold test (dashed-boxed in Figure 3-2) to decide whether a failure has occured. This is **Step 3** in our FDI procedure as outlined in Section 3.1. In section 3.4.1, we show that any failure, whether in a sensor or a valve, or in an internal component, will shift the mean of the residual. The description and design considerations of the preliminary threshold test is

48

$\hat{y}$ = predicted measurement
y = sensor measurement
e = preliminary threshold
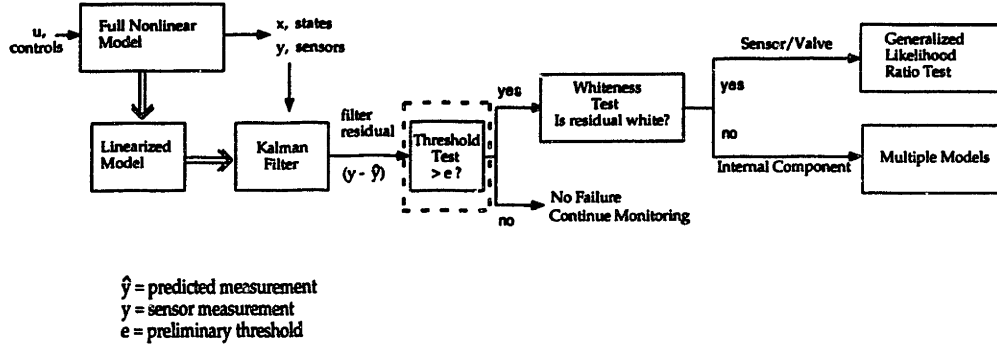
Figure 3-2: SSME VHM System

discussed in Section 3.4.2.

3.4.1 Failures and the Residual's Mean

In this section we demonstrate that whenever there is a failure of any kind, the mean of the residual will shift. Suppose that the true system is given by Eqs.(3.15-3.16) with the usual notations (repeated here for reference).

$$x(k+1) = Ax(k) + Bu(k) + \Gamma w(k) \qquad (3.32)$$

$$y(k) = Cx(k) + v(k) \qquad (3.33)$$

First, we want to show that the residual is biased when the failure occurs in a sensor or valve. Suppose that in the presence of a sensor failure, Eqs.(3.32-3.33) becomes

$$x(k+1) = Ax(k) + Bu(k) + \Gamma w(k) \qquad (3.34)$$

$$y(k) = Cx(k) + v(k) + f(k) \qquad (3.35)$$

where $f(k)$ represents the step failure with magnitude one. Substituting Eq.(3.35) into Eqs.(3.21,3.23-3.24) yields

$$\hat{x}^-(k+1) = A\hat{x}^-(k) + Bu(k) + AK\varrho(k) \qquad (3.36)$$

49

Using Eqs.(3.34-3.35), the error dynamics of this filter $\tilde{x}^-(k)$ can be expressed as

$$\tilde{x}^-(k+1) = x(k+1) - \hat{x}^-(k+1) \tag{3.37}$$

$$\tilde{x}^-(k+1) = (A - AKC)\tilde{x}^-(k) - AKv(k) + \Gamma w(k) - AKf(k) \tag{3.38}$$

and

$$E[\tilde{x}^-(k+1)] = (A - AKC)E[\tilde{x}^-(k)] - AKf(k) \tag{3.39}$$

Clearly, $\tilde{x}^-(k+1)$ is biased because of the term $f(k)$. The biased error dynamics in turn implies that the residual is biased since

$$E[\varrho(k+1)] = CE[\tilde{x}^-(k+1)] \tag{3.40}$$

Suppose next that the failure occurs in a control valve at time k:

$$x(k+1) = Ax(k) + Bu(k) + \Gamma w(k) + f(k) \tag{3.41}$$

$$y(k) = Cx(k) + v(k) \tag{3.42}$$

and again using Eq.(3.23) and Eq.(3.41), we can write

$$\hat{x}^-(k+1) = (A - AKC)\left\{A\hat{x}^-(k) + AKC\tilde{x}^-(k) + AKv(k)\right\}$$
$$+ AKCAx(k) + AKCBu(k) + AKCf(k) \tag{3.43}$$

$$x(k+1) = A^2x(k) + ABu(k) + Af(k) + Bu(k+1) + f(k+1) \tag{3.44}$$

These two equations yield the error dynamics

$$\tilde{x}^-(k+1) = \left(A^2 - \{(A - AKC)AKC + AKCA\}\right)\tilde{x}(k) + f(k+1)$$
$$(A - AKC)f(k) - (A - AKC)AKv(k) \tag{3.45}$$

which is also biased because of the terms $f(k)$ and $f(k+1)$.

Consider next the case of an internal component failure where the filter is based on a model with incorrect parameter values:

$$\hat{x}^-(k+1) = A_1 \hat{x}^-(k) + B_1 u(k) + A_1 K(k)(y(k) - C_1 \hat{x}^-(k)) \qquad (3.46)$$

where A_1, B_1, C_1 are different from A, B, C as a result of the failure. The error dynamics is then

$$\begin{aligned}\tilde{x}^-(k+1) &= (A_1 - A_1 K(k) C_1)\tilde{x}^-(k) + (\Delta A - A_1 K(k) \Delta C)x(k) \\ &\quad + \Delta B u(k) + \Gamma w(k) + A_1 K(k)v(k)\end{aligned} \qquad (3.47)$$

where $\Delta A = A - A_1$, $\Delta B = B - B_1$, and $\Delta C = C - C_1$. The error dynamics in Eq.(3.47) is biased because of the term $\Delta A, \Delta B, \Delta C$, and the mean of $x(k)$ itself. This is also true even if ΔB is zero since the mean of $x(k)$ is always nonzero. To see this, recall from Section 2.3 that $x(k)$ represents the difference between the actual state and the steady state. The actual state always deviates from the steady state when either ΔA or ΔC is nonzero or when both ΔA and ΔC are nonzero, which implies that the mean of $x(k)$ cannot be zero.

Finally, we present simulation experiments demonstrating the shift of the residual's mean in the presence of sensor/valve and internal component failures. Figure 3-3 shows the residual with a step failure in the *HPOTP rotational velocity* sensor occurs at time $t = 1$ second. The residual is unbiased prior to 1 second, and its mean shifts after the failure occurs. Similar results are demonstrated in Figure 3-4 and Figure 3-5 where the failures occur in a control valve (*fuel preburner oxidizer valve*) at time $t = 1$ second and in an internal component (bearing failure) at time $t = 0.8$ second, respectively.

3.4.2   Description and Design of the Threshold Test

The purpose of the threshold test is to detect the occurence of any failure. While both sensor/valve and internal component failures shift the mean of the

51
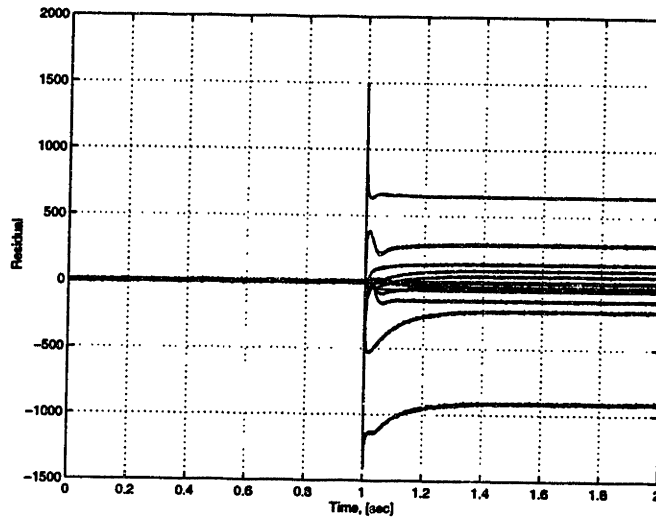
Figure 3-3: Residual of a HPOTP rotational speed sensor failure. Failure occurs at 1 second.
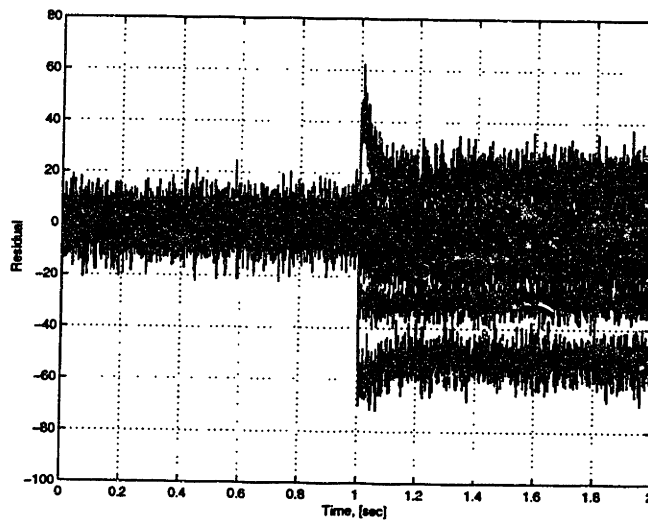
Figure 3-4: Residual of a main fuel valve failure. Failure occurs at 1 second.
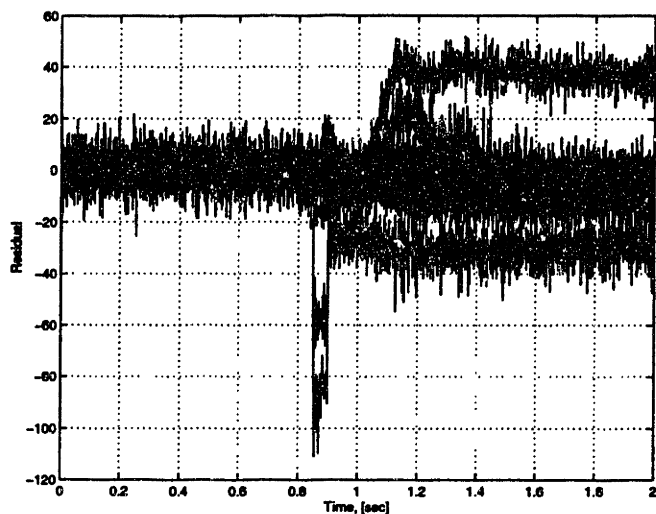
Figure 3-5: Residual of an internal component failure (HPOTP bearing). Failure occurs at 0.8 second.

residual, their effect on the covariance of the residual is different. As shown in [4], when the failure in a sensor or a valve is modeled as an additive input, the covariance of the residual is the same as it was before the failure occurs. The residual's covariance, however, is changed in the presence of internal component failures. It can only be numerically estimated from the sample data since an analytical form is not obtainable due to the lack of knowledge of the new system matrices (matrices A_1, B_1, and C_1 as shown in the discussion of internal component failures in the previous section). Since the type of the anticipated failure is generally unknown, a good detection test must take into account the lack of information on the residual's covariance. A test commonly used for this particular situation is the t-test which uses the sample covariance. The procedure of this test is outlined below with a few numerical examples. A thorough treatment of this test can be found in [2].

From the discussions in the previous sections, knowing that the residual is Gaussian with zero-mean and known covariance in the absence of failure, and otherwise Gaussian with mean m and unknown covariance, we consider the

following hypotheses for the interval $k_0, ..., k_0 + N$:

$$H_0 : m = 0$$
$$H_1 : m \neq 0 \qquad (3.48)$$

where $m = \{m(k_0), ..., m(k_0 + N)\}$. As shown in [2], the standard procedure for testing hypotheses (3.48) for N samples of the observed residual using the t-test is to define the T^2 statistics

$$T^2 = N\bar{\varrho}'S^{-1}\bar{\varrho} \qquad (3.49)$$

which has t-distribution with $N - 1$ degrees of freedom. The notation $\bar{\varrho}$ denotes the sample mean vector of the residual and S denotes the sample covariance matrix. These two quantities can be computed as

$$\bar{\varrho} = \frac{1}{N} \sum_{i=k_0}^{i=k_0+N} \varrho_i \qquad (3.50)$$
$$S = \frac{1}{N-1} \sum_{i=k_0}^{i=k_0+N} (\varrho_i - \bar{\varrho})(\varrho_i - \bar{\varrho})' \qquad (3.51)$$

and hypothesis H_0 is rejected if $T^2 \geq e$. The threshold, e, is chosen such that the probability of error (or the probability that $T^2 \geq e$ when H_0 is true) is equal to the significance level α_o. The value of α_o is a design choice. Once α_o is specified, e can be found using the t-distribution table.

To demonstrate this test, we present three numerical simulations. In Figures 3-6, 3-7, 3-8, the value of T^2 is computed using a sliding window of $N = 50$ points. The residual of these failures are illustrated in Figures 3-3, 3-4, and 3-5 respectively. Immediately after the failure, the value of T^2 jumps, indicating the presence of the failure. These failures are detected with a threshold $e = 40$. Since the noise is statistical, noise samples whose values are large can occasionally cause T^2 to cross the threshold, thus triggering a false alarm. On the other hand, raising the threshold may delay detection of true failures. Error due to
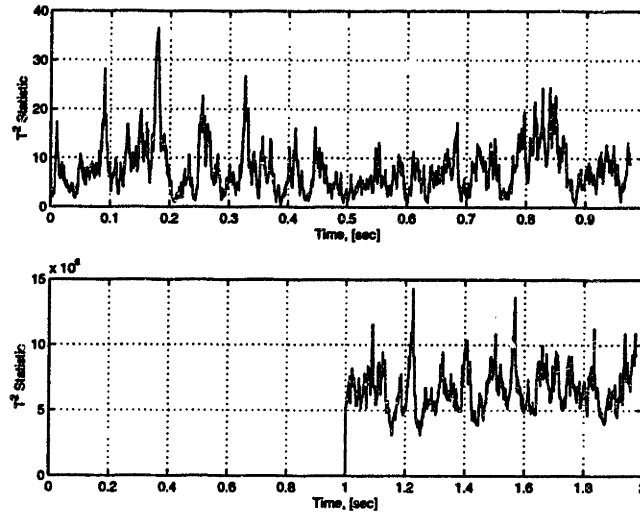
Figure 3-6: T^2 Statistic for the residual in the absence of failure (above) and in the presence of failure (below). The failure occurs in the HPOTP rotional speed sensor at 1 second.

linearization must also be taken into account when determining the threshold level. The threshold level therefore represents a tradeoff between false alarm rate and detection speed. For instance, if the threshold is lowered to $e = 30$, the failure would be detected slightly earlier but at the expense of possibly causing false alarm at time $t = 0.17$ seconds.

We want to emphasize that these failures can also be detected by just observing the residual, i.e. eye-balling or calculating the residual's mean. This is the case when the jump in the residual is large and obvious. In contrast, if the jump in the residual is small and is contaminated with noise, deciding whether a failure has occured can be quite subtle. The presence of noise implies that in general the detection is not always perfect. However, we want a detection test as good as possible in an appropriate sense.
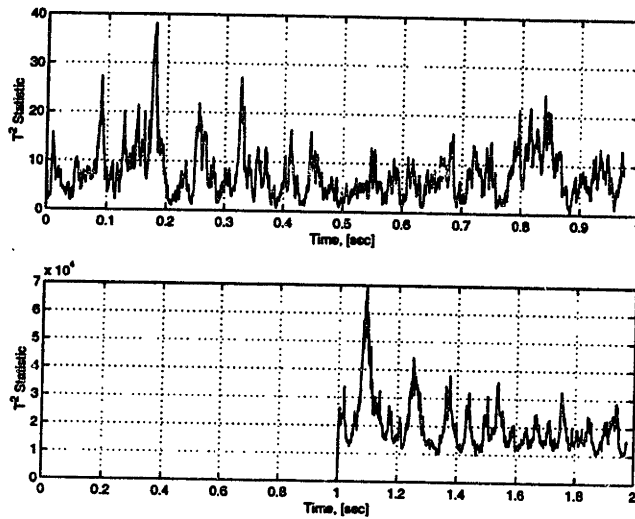
Figure 3-7: T^2 Statistic for the residual in the absence of failure (above) and in the presence of failure (below). The failure occurs in the main fuel valve at 1 second.
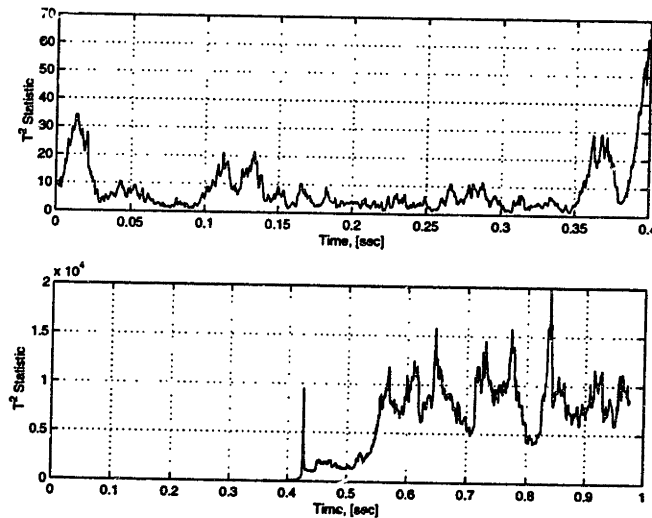
Figure 3-8: T^2 Statistic for the residual in the absence of failure (above) and in the presence of failure (below). The failure occurs in an internal component (HPOTP bearing) at .4 seconds.

$\hat{y}$ = predicted measurement
y = sensor measurement
e = preliminary threshold
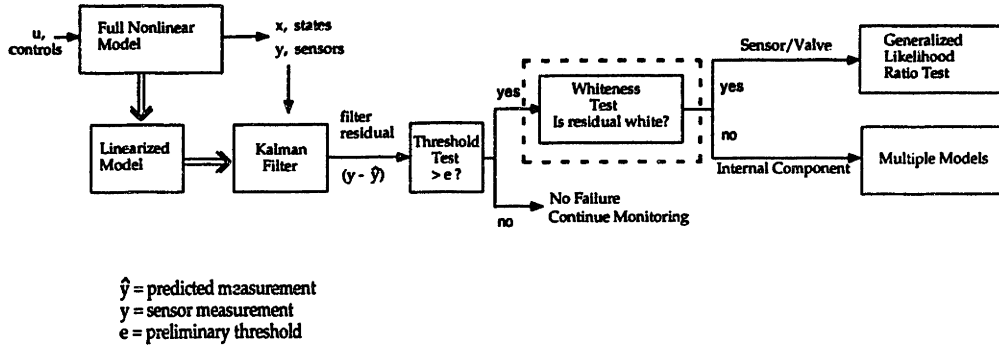
Figure 3-9: SSME VHM System

3.5 The Whiteness Test

Once the failure is detected, we proceed to the first level of isolation. This level categorizes the failure as either a sensor/valve failure or an internal component failure based on the whiteness of the residual. This is **Step 4** in our outline (dashed-box in Figure 3-9). We first show in Section 3.5.1 that when the failure occurs in a sensor or a valve, the residual is white, and when the failure is in an internal component, the residual is not white. In Section 3.5.2, the design and description of the whiteness test is discussed along with some numerical simulation results.

3.5.1 Whiteness of the Residual

Let us first show that the residual is white when the failure occurs in a sensor or a valve. Let $\varrho_0(k)$ and $\varrho_1(k)$ denote the residuals in the absence and presence of failure, respectively. Due to the linear feature in the dynamic equations and filter equations, we can decompose the residual $\varrho_1(k)$ into

$$\varrho_1(k) = \varrho_0(k) + g(k) \tag{3.52}$$

where $g(k)$ is a deterministic but unknown quantity representing the effect of the failure. In order to show that $\varrho_1(k)$ is white, we must show that the auto

covariance sequence $K_{\varrho_1 \varrho_1}(k,n)$ is zero for $n \neq k$. That is,

$$K_{\varrho_1 \varrho_1}(k,n) = E\{\varrho_1(k)\varrho_1(n)'\} - E\{\varrho_1(k)\}E\{\varrho_1(n)\} = 0 \qquad (3.53)$$

for $n \neq k$. Substitute Eq.(3.52) into the above equation and rearrange yields

$$
\begin{aligned}
K_{\varrho_1 \varrho_1}(k,n) &= E\{\varrho_0(k)\varrho_0(n)'\} + E\{\varrho_0(k)\}g(n)' \\
&\quad + g(k)E\{\varrho_0(n)'\} + E\{g(k)g(n)'\} - E\{g(k)g(n)'\} \\
&= 0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (3.54)
\end{aligned}
$$

where we have used the uncorrelation property of the residual from Eq.(3.30) so that the first term is zero, and the zero-mean property from Eq.(3.29) for the second and third term to be zero. Hence, the whiteness property of the residual is preserved when the failure occurs in a sensor or a valve.

We now show that when the residual is not white, the failure is assumed to have occured in an internal component. Before we present our argument, we first make some comments about white noise process and the optimality of the Kalman filter. Generally, white noise can perhaps be seen as the most unpredictable stochastic process. Hence, if the residual of the Kalman filter is white, then this means that the filter already utilized all the available information, and therefore it must be optimal. The optimality here is in the sense of generating the conditional mean (which the KF does). From Section 3.3, we showed that the optimality of the filter is a sufficient condition for the whiteness of the residual. Hence, if the residual is not white, then one can conclude that the filter is suboptimal. The non-whiteness of the residual is usually a consequence of using the wrong model to generate the residual. That is, the plant used in the realization of the Kalman filter is not the same as the actual plant. Whenever an internal component failure occurs in the system, the actual plant deviates from the nominal plant on which the Kalman filter is based, causing the filter to become suboptimal. While it is true that there are other changes in the system

besides internal component failures that account for the suboptimality, from our experience these changes generally appear to occur so infrequently that we can ignore them. In summary, if the residual is not white, then it is assumed that the failure occurs in an internal component.

3.5.2 Description and Design of the Whiteness Test

The whiteness test is introduced in this section to test the whiteness of the residual. Specifically, as shown in [13, 17], the test checks the correlation of each component of the the residual at different time instants. The first step in this test is to estimate the auto-correlation function (acf), $c_{\rho\rho}(k)$, and the normalized acf, $r_{\rho\rho}$, of all components of the residual as follows

$$c_{\rho\rho}(k) = \frac{1}{N} \sum_{t=k_0}^{t=N-k} (\rho_t - \bar{\rho})(\rho_{t+k} - \bar{\rho}) \tag{3.55}$$

$$r_{\rho\rho}(k) = \frac{c_{\rho\rho}(k)}{c_{\rho\rho}(0)} \tag{3.56}$$

for k = 1,2,...N-1. The notation ρ denotes an arbitrary component of the residual vector ϱ. The mean of this component, $\bar{\rho}$, is given by

$$\bar{\rho} = \frac{1}{N} \sum_{i=k_0}^{i=k_0+N} \rho_i \tag{3.57}$$

It is shown in [13] that if the residual is indeed white and the number of terms in the residual sequences, N, is large, then $r_{\rho\rho}(k)$ follows the statistics of a Gaussian random variable with zero mean and variance 1/N (these statistics are usually denoted as $\mathcal{N}(0, \frac{1}{N})$). One common method to check the statistics of $r_{\rho\rho}(k)$ is to specify a significance level of the test at 5 percent. That is, $r_{\rho\rho}(k)$ must lie in the region $\pm 1.96/\sqrt{N}$ more than 95 per cent of the times for the residual to be white. The significance level is essentially the probability that establishes a significant enough agreement from expected distribution to

justify that the hypothesis that $r_{\rho\rho}(k)$ is $\mathcal{N}(0, \frac{1}{N})$. We want to note that in this test, the cross auto-correlation function of different components of the residual at different time lags can also be incorporated into this test using the same procedure, but for simplicity it is not included in this test.

Recall from chapter 2 that the residual vector has fifteen components. In the implementation of this test, the decision rule is that if about 13 out of 15 acf's are inside the significance interval 95 percent of the time, then the residual is declared white. The choice of 13 acf's depends on several considerations and can be different based on the information available. For example, our choice of 13 acf's is based on our consideration of the number of data points N used to compute the acfs and the level of contamination in the output signal due to sensor noise and process noise. In practice, with the actual measured data, one should also take into consideration modeling errors that may cause the residual to be correlated when it it actually white. One major modeling error that usually arises is the inconsistency between the mathematical model of the engine and the actual engine itself.

The application of the above test is illustrated in Figures 3-10 and 3-11 for sensor/valve failures, and Figure 3-12 for an internal component failure. Figure 3-10 plots the normalized acf, $r_{\rho\rho}$, of each component of the residual before and after the failure for an interval of N = 1000 data points. The two solid horizontal lines on each plot represent the region $\pm 1.96/\sqrt{N}$. The failure is a bias in the *HPOTP rotational velocity* sensor. These plots show that the values of the acf of each component of the residual after the failure lies inside the significance level threshold (the two horizontal solid lines) more than 95 per cent of the times. The acf of each component of the residual before the failure also shows the same result as expected. Similar conclusions are obtained for the failure in one of the actuators (*fuel preburner oxidizer valve*) as illustrated in Figure 3-11. However, the acf's are different for internal component failures. Figure 3-12 is a typical acf plot when the failure occurs in an internal component (*bearing failure and polytropic efficiency degradation in HPFTP*). About 9 out of 15

60
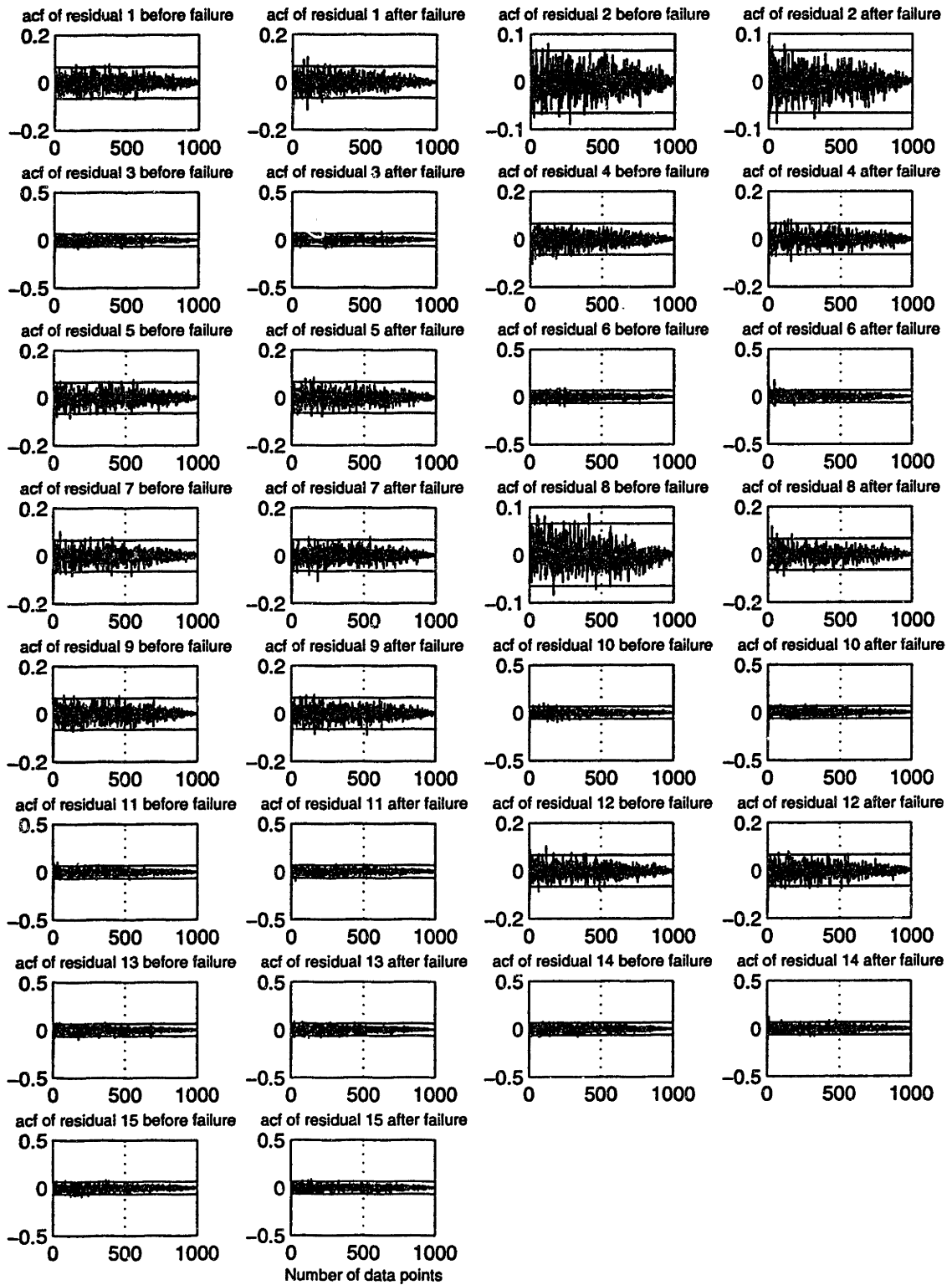
Figure 3-10: acf of a sensor-failure residual. Residuals 1-15 are all white.
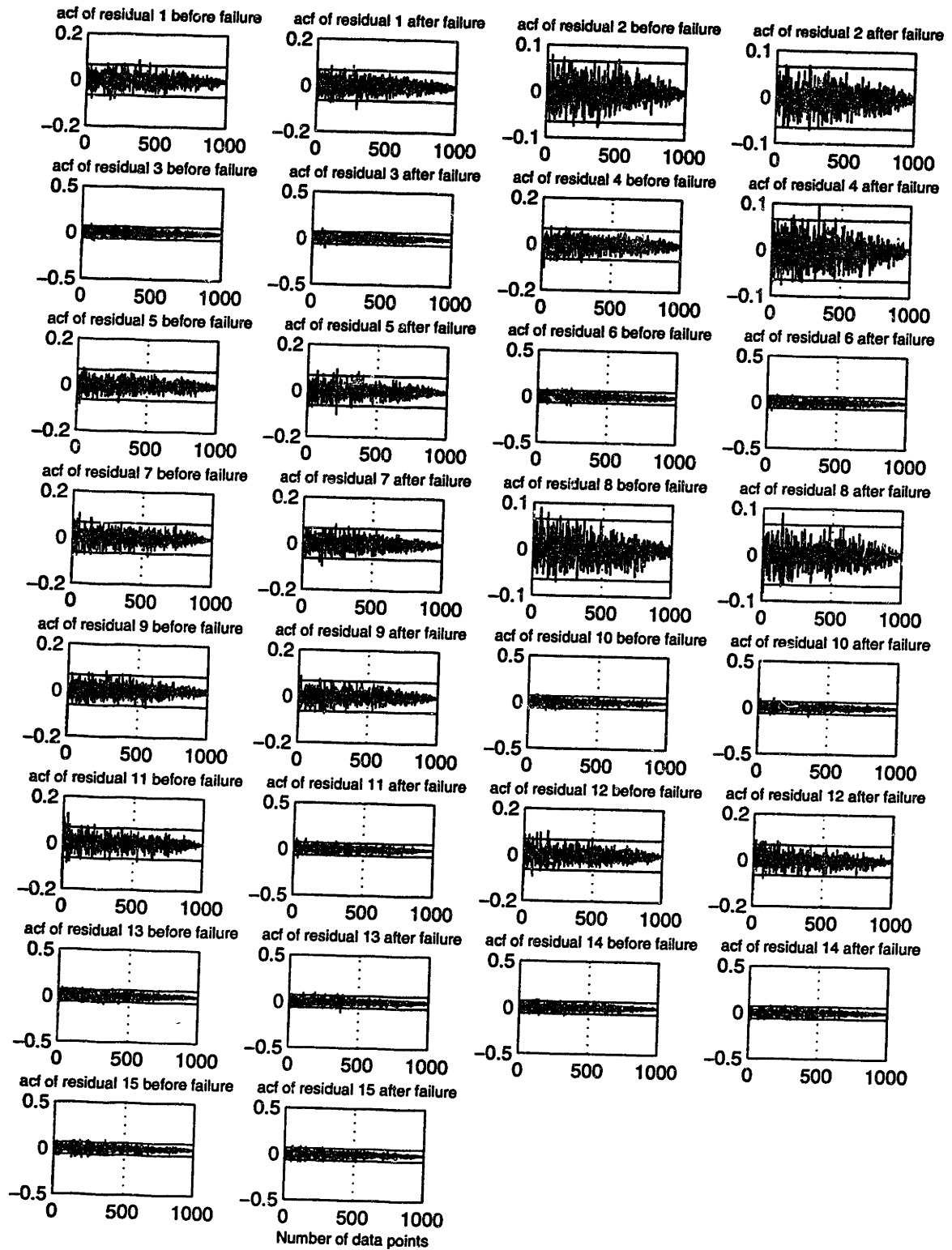
Figure 3-11: acf of a valve-failure residual. Residuals 1-15 are all white.
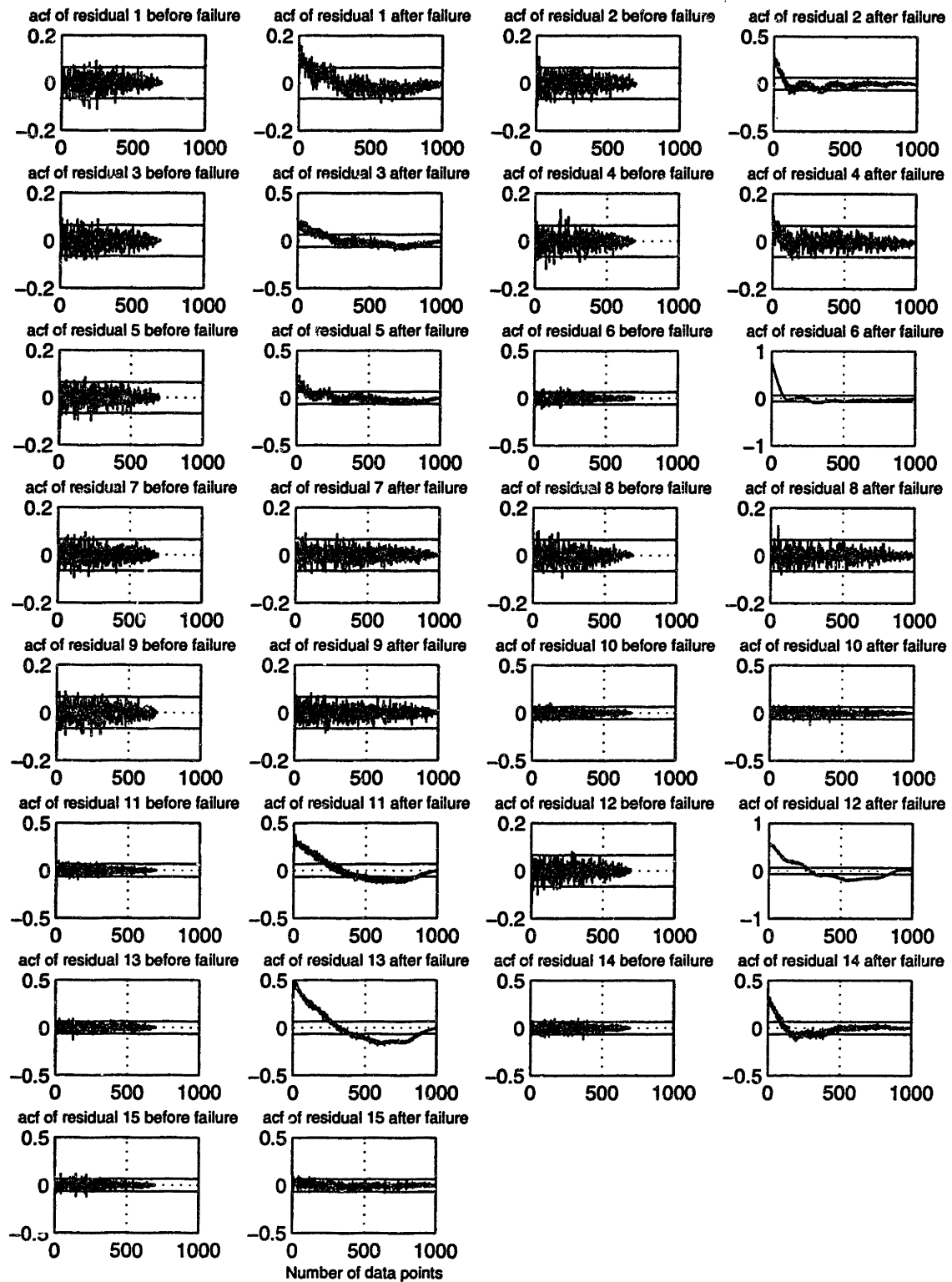
Figure 3-12: acf of an internal component-failure residual. After the failure, residuals 4,7,8,9,10,15 are white while residuals 1,2,3,5,6,11,12,13,14 are not white.

acf's lie outside the significance interval comfirming that the residual of internal component failures is not white.

Chapter 4

FDI for Sensor and Valve Failures

In this chapter, we use the linearized state space model described in Section 2.3 to design the Generalized Likelihood Ratio Test (GLRT) for the detection and isolation of failures in the engine control valves and sensors. As indicated in Figure 4-1, these failures, as well as any others, are first detected by the threshold test. Then the whiteness test determines whether the whiteness of the residual has changed. As explained in Chapter 3, if the residual is white, then it can be assumed that the failures occured in either a sensor or a valve. The GLRT (dashed-box in Figure 4-1) is then used to isolate the specific sensor

$\hat{y}$ = predicted measurement
y = sensor measurement
e = preliminary threshold
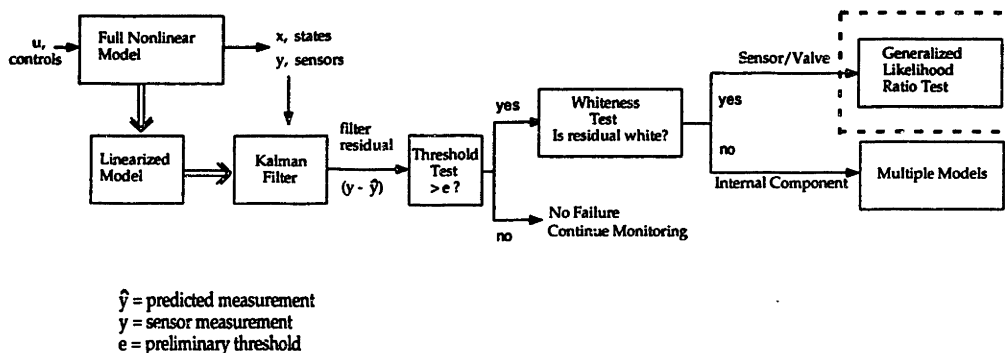
Figure 4-1: SSME Vehicle Health Monitoring System

or valve that failed.

The chapter is organized as follows. Section 4.1 describes the role of the Kalman filter in the GLRT for failure detection and isolation. Specifically, using the the observed measurements along with the SSME mathematical model, a KF is designed to generate the residual. The GLRT uses a statistic that is a function of the the residual to identify the faulty sensor or valve. In Section 4.2, simulation results are presented for failures in the engine's fifteen sensors and five valves. A geometrical interpretation of the GLRT is presented in Section 4.3 where the issue of the distinguishablitity of one failure from another is discussed.

4.1   The GLRT

We describe here a simple version of the GLRT. Specifically, we consider only step failures, and we assume that the failure either exists or does not exist for the interval of interest. Nevertheless, we demonstrate that our method works for more general failures. More general versions of the GLRT that consider a larger class of failure modes, or that takes into account the failure time, are derived in [19, 30]. The GLRT described below is for linear Gaussian systems, but can also be applied to nonlinear systems if the linearization error are insignificant as we demonstrate. The no failure (H_0) and failure (H_j) hypotheses are

$$H_0 \quad : \quad x(k+1) \quad = \quad Ax(k) + Bu(k) + \Gamma w(k) \tag{4.1}$$

$$y(k) \quad = \quad Cx(k) + v(k) \tag{4.2}$$

$$H_j \quad : \quad x(k+1) \quad = \quad Ax(k) + Bu(k) + \Gamma w(k) + F_j s_{k-\tau} \cdot \nu \tag{4.3}$$

$$y(k) \quad = \quad Cx(k) + v(k) + L_j s_{k-\tau} \cdot \nu \tag{4.4}$$

for $k\epsilon[k_o, k_f]$, and $j = 1, 2, ..., 20$ since we have five control valves and fifteen sensors. The vectors F_j and L_j describe the way the failure is injected into the system. For instance, if the failed component is a valve, then F_j is the column

66

vector of B representing that particular valve in the plant dynamics, and there is no L_j. The parameter τ^* is the unknown time at which the failure occurs, the function s_k is the unit step function representing the failure *mode*, and ν is the magnitude of the step failure, which is generally not known.

For observations over a finite interval $k_0 \leq k \leq k_f$, the generalized likelihood ratio of the above hypothesis test is given by

$$\Lambda(k_f, \hat{\nu}, j) = \frac{p(y(k_0), ..., y(k_f) \mid H_j, \hat{\nu})}{p(y(k_0), ..., y(k_f) \mid H_0)} \tag{4.5}$$

where $\hat{\nu}$ is the maximum likelihood estimate of ν assuming that hypothesis H_j is true.

$$\hat{\nu} = \arg \max_{\nu} p(y(k_0), ..., y(k_f)|H_j, \nu) \tag{4.6}$$

Note that τ^* is not included in the hypothesis test because of the assumption that the failure either exists or does not exist in the entire interval of interest $k_0 \leq k \leq k_f$. That is, in hypothesis H_j, $\tau^* = k_0$.

As motivated in Chapter 3, the above probability distributions are difficult to compute because the sequence $\{y(k_0), ..., y(k_f)\}$ is not independent. Therefore we replace that sequence by its corresponding residual or innovation sequence $\varrho(k)$

$$\varrho(k) \equiv y(k) - E(y(k) \mid y(0), ..., y(k-1)) \tag{4.7}$$

The $\varrho(k)$'s form an independent sequence. Note that the observation sequence $y(k)$ can be reconstructed from the residual sequence $\varrho(k)$, and vice-versa. Hence, both of these sequences contain the same statistical information. In terms of the residual, the hypothesis test of Eqs.(4.1-4.4) becomes

$$H_0 \quad : \quad \varrho(k) = \varrho_0(k) \tag{4.8}$$

$$H_j \quad : \quad \varrho(k) = \varrho_0(k) + G(j, k)\nu \tag{4.9}$$

where $G(j, k)$, the projection of the failure onto the residual output, can be

computed recursively as shown below. The two hypothesis tests of Eqs.(4.1-4.4) and Eqs.(4.8-4.9) are therefore equivalent.

To generate the residual and calculate $G(j,k)$, we implement a Kalman filter based on the no-failure model from Eqs.(4.1-4.2). These equations are the same as the ones presented in Section 3.3:

$$\hat{x}^-(k+1) = A\hat{x}^+(k) + Bu(k) \tag{4.10}$$

$$\hat{x}^+(k+1) = \hat{x}^-(k+1) + K(k+1)\varrho(k+1) \tag{4.11}$$

$$\varrho(k+1) = y(k+1) - C\hat{x}^-(k+1) \tag{4.12}$$

$$P^-(k+1) = AP^+(k)A' + \Gamma Q\Gamma' \tag{4.13}$$

$$V(k) = CP^-(k)C'(k) + R \tag{4.14}$$

$$K(k) = P^-(k)C'(k)V(k)^{-1} \tag{4.15}$$

$$P^+(k) = P^-(k-1) - K(k)CP^-(k-1) \tag{4.16}$$

Because of the linearity in the state equations and the filter equations, we can express

$$x(k) = x_0(k) + x_j(k)\nu \tag{4.17}$$

$$\hat{x}^-(k) = \hat{x}_0^-(k) + \hat{x}_j^-(k)\nu \tag{4.18}$$

$$\hat{x}^+(k) = \hat{x}_0^+(k) + \hat{x}_j^+(k)\nu \tag{4.19}$$

$$\varrho(k) = \varrho_0(k) + G(j,k)\nu \tag{4.20}$$

where $x_0(k)$, $\hat{x}_0^-(k)$, $\hat{x}_0^+(k)$, and $\varrho_0(k)$ are the responses in the absence of failure, while x_j, $\hat{x}_j^-(k)$, $\hat{x}_j^+(k)$, and $G(j,k)$ are responses that only appear in the presence of failure in component j. These terms can be recursively written as follows:

$$x_j(k+1) = Ax_j(k) + F_j\nu \tag{4.21}$$

$$\hat{x}_j^+(k+1) = \hat{x}_j^-(k+1) + K(k+1)G(j,k+1) \tag{4.22}$$

$$\hat{x}_j^-(k+1) = A\hat{x}_j^+(k) \tag{4.23}$$

Let $\Upsilon(j,k) = x_j(k) - \hat{x}_j^-(k)$, it follows that

$$\Upsilon(j,k+1) = A(I - K(k)C)\Upsilon(j,k) + [F_j - K(k)L_j] \qquad (4.24)$$

$$G(j,k) = C\Upsilon(j,k) + L_j \qquad (4.25)$$

where $G(j,k)$, the failure signature, is simply the output of the dynamics of $\Upsilon(j,k)$.

Using the residual and the failure signature, the likelihood ratio of Eq.(4.5) is given by

$$
\begin{aligned}
\Lambda(k_f,\nu,j) &= \frac{p(\varrho(k_0),...,\varrho(k_f) \mid H_j,\nu)}{p(\varrho(k_0),...,\varrho(k_f) \mid H_0)} \\
&= \prod_{k=k_0}^{k=k_f} \frac{p(\varrho(k) \mid H_j,\nu)}{p(\varrho(k) \mid H_0)} \qquad (4.26)
\end{aligned}
$$

where the last equality follows from the independence of the residual process. By taking the log of the above equation, we obtain the following simple expression for the likelihood ratio in terms of the residual

$$
\begin{aligned}
\ell(k_f,\nu,j) &\equiv \log \Lambda(k_f,\nu,j) \\
&= \nu\chi(k_f,j) - \frac{1}{2}\nu^2 S(k_f,j) \qquad (4.27)
\end{aligned}
$$

where $\chi(k_f,j)$ is a weighted linear combination of the residuals representing the correlations of the residuals and the failure signature $G(j,k)$. Specifically,

$$\chi(k_f,j) = \sum_{k=k_0}^{k=k_f} G'(j,k)V(k)^{-1}\varrho_j(k) \qquad (4.28)$$

and $S(k_f,j)$ is a known quantitiy

$$S(k_f,j) = \sum_{k=k_0}^{k=k_f} G'(j,k)V(k)^{-1}G(j,k) \qquad (4.29)$$

The *generalized* likelihood ratio is given by

$$\ell(k_f, j) = \max_{\nu} \ell(k_f, \nu, j) \qquad (4.30)$$

The failure magnitude ν is obtained by computing its maximum likelihood estimate

$$\hat{\nu}(k) = \frac{\chi(k_f, j)}{S(k_f, j)} \qquad (4.31)$$

Substitute Eq.(4.31) into Eq.(4.30) gives

$$\ell(k_f, j) = \frac{\chi(k_f, j)^2}{S(k_f, j)} \qquad (4.32)$$

The on-line implementation of the GLRT consists of computing the quantity $S(k_f, j)$ and $\chi(k_f, j)$ for a time interval of interest $k_0 \leq k \leq k_f$ and comparing the likelihood ratio $\ell(k_f, j)$ against a selected threshold for failure detection. This detection test is essentially identical to the threshold test presented in Chapter 3. However, the GLRT algorithm requires the computation of each $\ell(k_f, j)$ whereas the threshold test only requires a single computation of the T^2 statistics using Eq.(3.49). Threshold test is therefore used for detection to reduce the computation complexity.

To isolate the faulty component, we note that each hypothesis test in Eqs.(4.8-4.9) produces a likelihood function $\ell(k_f, j)$. The most likely component to have failed, j^*, is the one that produces the maximum value of the likelihood ratio, i.e.

$$j^* = \arg \max_{j} \ell(k_f, j) \qquad (4.33)$$

for $j = 1, 2, ..., 20$.

4.2 Results

In this section we present some numerical results for the test described in the previous section using the SSME model developed in Section 2.2. The Kalman filter is first implemented to track the propagation of the state variables. Figure 4-2 shows the filter state estimates (dot) and the actual value of the states (solid) generated from the nonlinear Rocketdyne model. These plots show that the filter can track the value of the engine variables very well. The insignificant error arises from the fact that the filter design assumes a linear plant while the measurement data input into the system originates from a nonlinear plant. The knowledge of the engine variables is valuable for heath monitoring because these estimates can be part of the flight's health records, and can be used for comparison.

The residual generated by the filter is then used in the GLRT algorithm for failure detection and isolation. Before discussing failure isolation, we illustrate how the detection test (of the GLRT) works. In a previous work [18], we applied the test to the HPOTP subsystem of the SSME where we used only three sensors. The time history of the three output residuals, averaged using a sliding window of length $k_f - k_0 + 1$ ($= 50$), is plotted in Figure 4-3. The units in Figure 4-3 are not shown because the three sensors measure three different quantities. A failure is introduced at time t= .1 seconds and the magnitude of the failure is approximately half of its steady state value. For a linear system, the residual would be unbiased. The small bias in one of the residuals prior to the failure is due to linearization error. Immediately after the failure, two of the residuals experience a jump in value.

For the entire SSME with fifteen sensors and five control valves, the residual plot is shown in Figure 4-4, where a failure in the *mass flowrate through the oxygen pump boost stage sensor* (Sensor #2) is chosen as a representative example. The failure is injected at time t = 1 second and the magnitude of the failure is approximately half of its steady state value. The likelihood ratio expression,
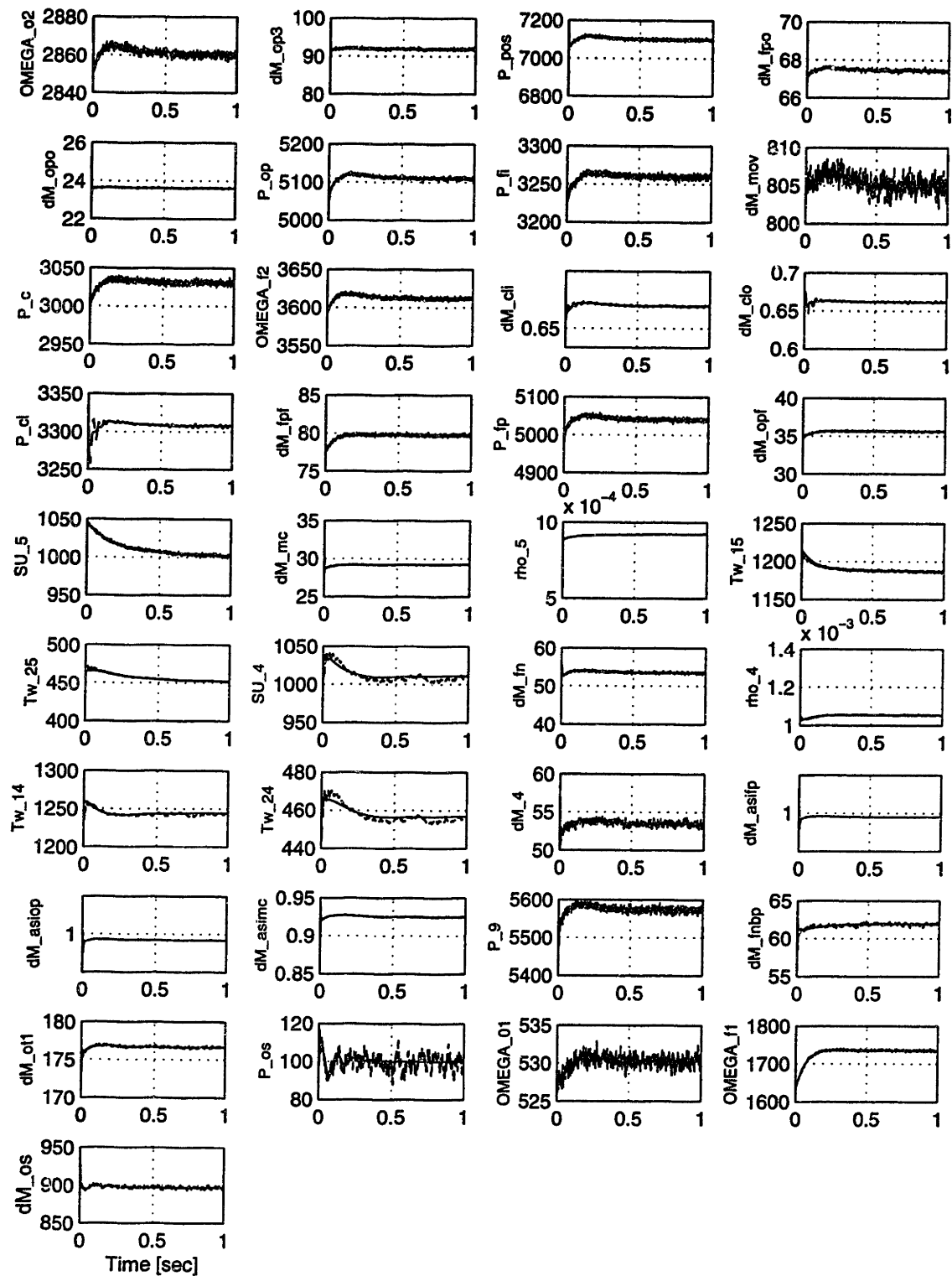
Figure 4-2: Kalman filter estimates (dot line) and actual states (solid line).
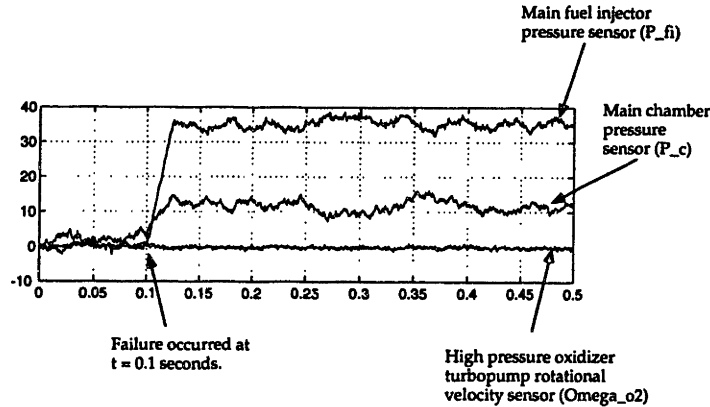
Figure 4-3: A failure's effect on the Kalman filter's residuals using the HPOTP model with 3 sensors.

Eq.(4.32), which is a function of the residuals and the hypothesized failures, indicates a jump as shown in Figure 4-5 for each specific hypothesized failure i.e. failure in a particular sensor or valve. This is an indication of an abrupt change, or a failure. In practice, a threshold is set, and a failure is declared once the likelihood ratio crosses it. Since the noise is statistical, noise samples whose values are large can occasionally cause the likelihood ratio expression to cross the threshold, thus triggering a false alarm. On the other hand, raising the threshold may delay detection of true failures. Error due to linearization must also be taken into account when determining the threshold level. The threshold level therefore represents a tradeoff between false alarm rate and detection speed.

Let us now discuss the isolation of the failure (in Sensor #2), presented above in Figure 4-4, which occurs at time equals 1 second. Figure 4-6 shows the likelihood ratio of each hypothesized failure on the same y-axis scale. These plots shows that the likelihood ratio of the true failure (in Sensor #2) is larger than the others, implying correct isolation of the faulty sensor and confirming Eq.(4.33). Failures in the other nineteen components were also successfully diagnosed with this procudure.

73
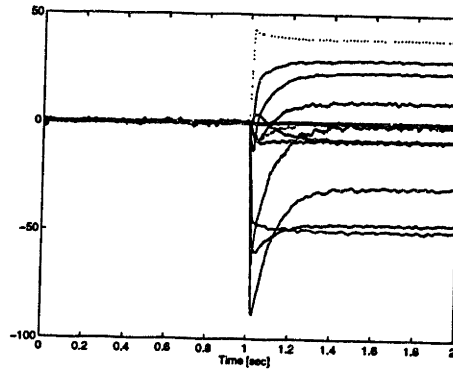
Figure 4-4: A failure's effect on the Kalman filter's residual using the complete model. The failure is in Sensor #2.
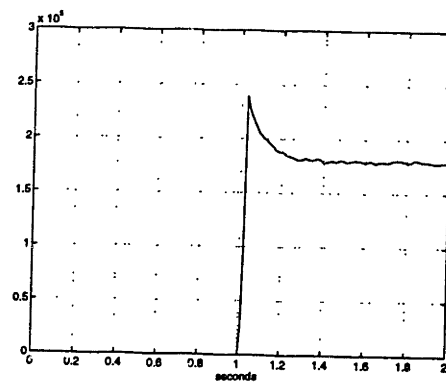
Figure 4-5: The Likelihood Ratio for Figure 4-4.
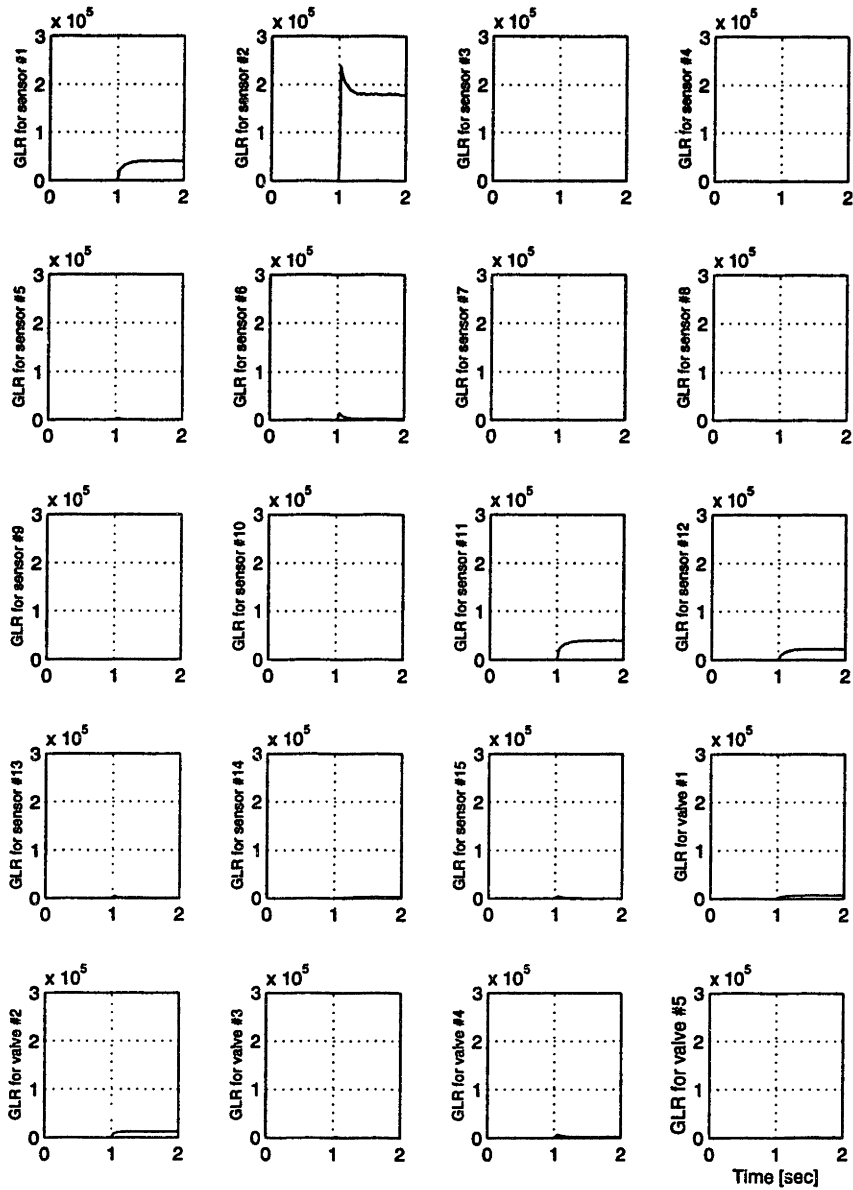
Figure 4-6: The Likelihood Ratios of the multiple hypothesis test for Figure 4-4. Failure occurs at t = 1 second.

4.3 Geometrical Intepretation

More insight can be gained into the above failure isolation procedure by looking at a geometrical interpretation. This interpretation relies on the fact that for each failure, the residual vector takes a unique and fixed direction in the residual space. For example, consider a sample of the residual of Figure 4-3 at any time after a failure occurs [18]. This sample is a vector in three dimensional space. The direction of this vector is an indicator of the failed component. Specifically, Figure 4-7 shows residual samples of several different failures for the HPOTP subsystem. Samples from the same failure are clustered together, and when averaged, they define a one-dimensional subspace(direction) that is identified with the failure. Different failures are therefore identified with different subspaces or directions. Each subspace is a line passing through the origin. Since, in our work, all failure modes are modeled as step functions, the resulting failure subspaces are simply straight lines. Generally, if the failure modes are arbitrary, then these subspaces become larger and more complex.

Given samples of an output residual, consider the projection of these samples, or their average, onto all the subspaces representing the failures under consideration. The projection with the largest magnitude indicates the failed component. Specifically, the failure signature $G(j,k)$ is the vector spanning the subspace of failure j at time k. The likelihood ratio $\ell(k_f,j)$ can be rewritten as

$$\ell(k_f,j) = \frac{1}{S(k_f,j)} \left(\sum_{k=k_0}^{k=k_f} < (R^{-1}\varrho(k))', G(j,k) > \right)^2 \qquad (4.34)$$

Each element in the above summation is a *weighted* projection of the residual onto the subspace spanned by $G(j,k)$. The summation attenuates the effect of noise.

The question now arises as to how distinguishable are any two failures. In steady state, $G(j,k) \to G(j)$ for all failures $j's$. The angles between the subspaces can define a distance measure. The larger the angle ranging from 0

76

Failure signatures of various failures:
1. Failure in fuel turbine flowrate
2.,3.,4. Sensor Failures (sensors 1,2,&3 respectively)
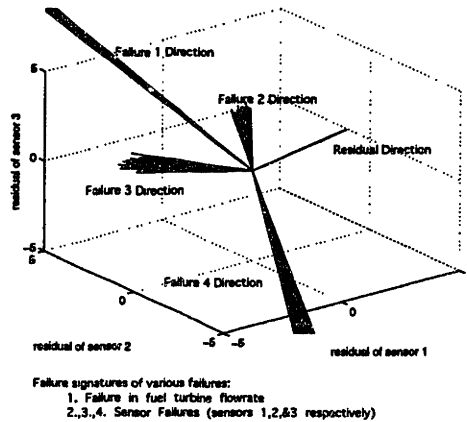
Figure 4-7: Residual vector samples from different failures for the HPOTP sub-system with 3 sensors.

to 90 degrees, the more distinguishable are the failures. Figure 4-8 illustrates the distinguishability of a failure in the *main fuel valve* (MFV, valve #4). The signature of the failure in the sensor (#15) measuring the *rotational speed of the HPFTP* Ω_{f2} is closest to that of the MFV's failure signature. This is not surprising if we consider the relationship between the MFV and Ω_{f2} sensor. A failure in either the MFV or the Ω_{f2} sensor would produces similar effect on other residuals, since these components regulate or indicate how much fuel is flowing through the rest of the engine. This kind of physical interpretation gives additional insight about the degree of distinguishability of various failures and can be employed at the design stage to place sensors judiciously.

For our complete model with fifteen sensors and five valves, Table 4.1 lists the angles between the residual vector of the failure in component j and that of the other nineteen failure signatures where each signature corresponds to a particular failure in component i. For instance, if the faulty component is the MFV (valve #4), then all the rows under the "valve # 4" column are angles between the *valve #4 failure* residual and the other signatures. The angles shown here may vary slightly depending on the level of noise. In our implementation, the variance of the process noise and sensor noise for a particular state or sensor is assumed to be about one percent of the steady state value. The angles shown
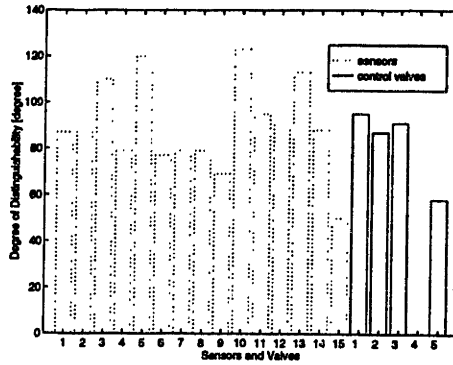
Figure 4-8: Distinguishability of Failure Signatures in MFV (valve #4).

in this table imply that failures in the twenty components are distinguishable, which is consistent with the results obtained by using the isolation procedure in Eq.(4.33).

The above discussion is concerned with failure in a single component. For simultaneous failures, it can be shown [20] that the maximum number of simultaneous control valve and sensor failures that can be isolated cannot exceed the number of sensors used. Thus with fifteen sensors and five control valves, one must be able to exclude by other means five candidate components. This is often done by triplicating sensors and if, for instance, five of the fifteen sensors are triplicated, then a voting system can be used to detect failures in these sensors. Simultaneous failures in the remaining ten sensors and in all the control valves can then be uniquely determined using the described method.

		sensors #														valves #				
		1	2	3	4	5	6	7&8	9	10	11	12	13	14	15	1	2	3	4	5
sensors #	1	0	74	88	88	85	81	81	79	83	32	19	65	65	85	76	58	81	87	80
	2	74	0	84	86	70	81	85	77	72	66	74	87	80	72	69	80	73	88	66
	3	88	84	0	26	83	65	63	90	86	88	89	76	79	81	85	89	84	70	86
	4	88	86	26	0	87	80	69	75	88	90	88	89	82	84	83	90	75	79	87
	5	85	70	83	87	0	76	63	85	24	88	84	60	89	12	87	86	82	60	88
	6	81	81	65	80	76	0	88	71	78	89	70	48	51	78	56	73	71	77	73
	7&8	81	85	63	69	63	88	0	88	77	84	79	73	83	68	87	89	74	79	83
	9	79	77	90	75	85	71	88	0	85	75	88	60	76	83	88	89	77	69	87
	10	83	72	86	88	24	78	77	85	0	86	82	62	88	24	88	88	82	57	88
	11	32	66	88	90	88	89	84	75	86	0	30	80	77	88	75	58	78	85	83
	12	19	74	89	88	84	70	79	88	82	30	0	62	55	85	68	52	77	89	82
	13	65	87	76	89	60	48	73	60	62	80	62	0	40	60	89	57	71	67	86
	14	65	80	79	82	89	51	83	76	88	77	55	40	0	90	90	43	57	88	87
	15	85	72	81	84	12	78	68	83	24	88	85	60	90	0	87	86	84	50	85
valves #	1	76	69	85	88	87	56	87	88	88	75	68	89	90	87	0	89	87	85	80
	2	58	80	89	90	86	73	89	89	88	58	52	57	43	86	89	0	89	87	86
	3	81	73	84	75	82	71	74	77	82	78	77	71	57	84	87	89	0	39	83
	4	87	88	70	79	60	77	79	69	57	85	89	67	88	50	85	87	89	0	58
	5	80	66	86	87	88	73	83	87	88	83	82	86	87	85	80	86	83	58	0

Table 4.1: Distinguishability of failure signatures for comple model sensors and valves.

79

Chapter 5

FDI for Internal Component Failures

In the previous chapter, the GLRT methodology was presented for detecting failures in the engine sensors and valves. As we discussed, the main characteristic of sensor and valve failures is that they shift the mean of the residual while preserving its whiteness property. For failures that take place in engine internal components, the mean of the residual shifts and the residual is colored.

In this chapter, the Multiple Model(MM) method is employed for detecting failures in the internal components of the engine. The premise here is that these failures have been detected by the threshold test and identified by the whiteness

$\hat{y}$ = predicted measurement
y = sensor measurement
e = preliminary threshold
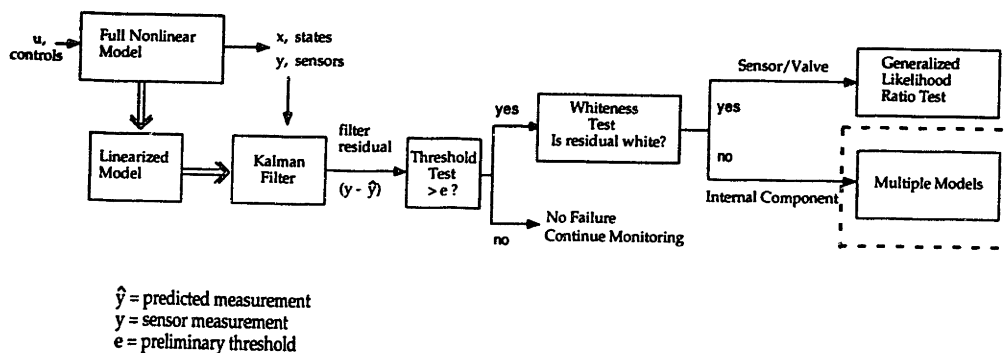
Figure 5-1: SSME Vehicle Health Monitoring System

test (Figure 5-1) as presented in Chapter 3. In Section 5.1, internal component failures are classified according to the subsystem in which they originate. Actual SSME failures that have occurred during past tests [8] are reviewed and categorized into different kinds, based on the rate at which they propagate throughout the engine. This categorization allows us to focus on the specific failures that are detectable by the MM technique. Section 5.2 provides the formulation of the MM structure, and compares and contrasts the MM technique with the GLRT. Simulation results from the implementation of the MM method are presented in Section 5.3 for a few representative failure scenarios described in [8]. In Section 5.4, the issue of robustness of the MM method to the choice of models is examined using an information theoretic distance criterion [3]. This distance suggests how to select models that are far apart enough so that the MM can distinguish between them.

5.1 Internal Component Failures

The internal component failures that we consider are those that occur in turbopumps, ducts, manifolds, injectors, nozzle, main combustion chamber, and many other components. In addition, these failures are ones that affect the performance of the engine. Many critical internal component failures that occurred in the past are documented in [8]. These failures can occur in numerous components in the engine and can cause substantial damage to neighboring and remote components, or to the entire engine as a whole, if they are not detected or compensated for.

Let us first consider the nature of these failures. The analysis of the failures compiled in [8] suggests that while failures that originate from different components progress at different rates, their degree of damage spreading to other components is unpredictable. This study also concludes that most failures typically occur at three different rates. Most turbopump and preburner failures had indications of trouble tens of seconds before becoming catastrophic. Main

81

burner and control failures developed at a slightly faster rate while duct, manifold, heat exchanger failures progressed in tenths of seconds. In our study, the focus is primarily on the failures at the first two rates since they develop at a slower pace then the failures of the third rate (tenths of seconds), and cannot easily be compensated. For instance, once there is a leakage in a duct, the leakage hole would be wide opened within the next few milliseconds and ruptured, since the pressure in the duct is extremely high.

Some of the internal component failures and their nature as reported in [8] are summarized in the following four examples. A full description of these failures as well as others can be found in [8] and references therein.

1. *Fuel Turbopump Failures:* High Pressure Fuel Turbopump failures are caused by a variety of factors such as turbine blade failures, cracks in the ducts, and coolant liner buckle. These failures usually result in high oxygen to fuel mixture ratio due to loss of fuel pumping. As a result, the gas temperature is raised above the design limits and damages the hot gas flow path.

2. *Oxidizer Turbopump Failures:* High Pressure Oxidizer Turbopump failures are due to design deficiencies in various forms such as bearing failures and LOX seal burning. These failures are typically followed by fire because of the addition of heat. In the case of bearing failure, rubbing is a common phenomenon that consequently added sufficient heat to cause fire, while in the case of LOX seal burning, the LOX and LH mixed to cause fire.

3. *Preburners and Main Burner Failures:* These failures are originated as a result of preburner LOX posts crack or main injector LOX posts crack due to design deficiencies (usually high cycle fatigue). The cracks in the posts allowed LOX to mix with hot gases and resulted in high mixture ratio and eventually fire.

4. *Duct and Manifolds Failures:* These failures occurred when the components of the nozzle such as ducts and manifolds ruptured. The ruptures were the results of low cycle fatigue and small cracks, that gradually became critical with time. These failures usually resulted in damage throughout the engine hot gas system.

These examples, along with the description of the engine presented in Chapter 2, provide a notion of the complex and tight interaction among the components of the SSME. The ultimate goal of any FDI algorithm is to identify the precise origin of the failure. However, there is no technique available that has this capability. As presented, the task of isolating the specific failed component, and how it results in further damage in other components, is a difficult task since this may require placing of sensors monitoring the specific components of interest. The space in the engine is extremely limited, and the environment is hostile (high temperature and pressure). For this reason, placing sensors in arbitrary locations is not practical. Besides the packaging constraint, adding extra sensors also implies extra costs. Due to these considerations, a more realistic approach is to determine which of the following subsystem has a faulty component: the High Pressure Fuel Turbopump, High Pressure Oxidizer Turbopump, the Low Pressure Oxidizer Turbopump, the Low Pressure Fuel Turbopump, the Main Injector, the nozzle, and the pipes/ducts. The term subsystem refers to a collection of components in the engine. For instance, the High Pressure Oxidizer Turbopump Subsystem is composed of the pump itself, the turbine, the injector, the oxidizer preburner, and other components therein.

Therefore, instead of attempting to detect and trace back the precise origin (component) of the failure, the VHM's algorithm specifies which subsystem exhibits anomalous behavior. For example, the algorithm identifies that the High Pressure Oxidizer Turbopump subsystem has failed, but not that a specific component(i.e. the turbine blade, or bearing) in that subsystem has malfunctioned. This approach is attractive and useful since internal component failures tend to quickly propagate from one component to another in the same subsystem be-

fore spreading to other subsystems. The knowledge of the degraded subsystem is crucial in guiding subsequent investigations such as using vibration monitoring to further analyze the identified faulty subsystem [24].

The internal component failures considered in this study can be modeled as changes in the geometry of the components in the engine. To represent these failures, we model them as changes in the parameter since the geometry is made of a set of parameters. For simplicity, no modeling of phenomenon such as fire is included. While it is not practical to model all parameters as they number about 150, we select a few in each subsystem or ones that are generally affected whenever there is a failure or an anomaly in that subsytem.

The following three examples illustrate the modeling of failures as parameter changes.

Example 1 : To model cracks in LOX posts as a 50 per cent change in the area of the in the main injector, the equation of the mass flow flow rate in the injector

$$\dot{m}_{fi} = B65 \frac{P_{fi}}{\sqrt{t_{fi}}} \sqrt{\left(\frac{P_c}{P_{fi}}\right)^{\frac{2}{1.4}} - \left(\frac{P_c}{P_{fi}}\right)^{\frac{2}{1.4}}} \tag{5.1}$$

is modified as

$$\dot{m}_{fi} = 0.5 B65 \frac{P_{fi}}{\sqrt{t_{fi}}} \sqrt{\left(\frac{P_c}{P_{fi}}\right)^{\frac{2}{1.4}} - \left(\frac{P_c}{P_{fi}}\right)^{\frac{2}{1.4}}} \tag{5.2}$$

where the parameter $B65$ represents the lumped area of the main injector, P_{fi}, P_c, and t_{fi} are the injector pressure, the main combustion chamber pressure, and the temperature in the the main injector, respectively. The 50 per cent change in the main injector area is modified as a decrease in 50 per cent in parameter $B65$.

Example 2 : Similarly, to model cracks in LOX posts as a 50 per cent decrease in the area of the fuel preburner injector of the High Pressure Fuel Turbopump subsystem, the equation of the mass flow rate into fuel preburner

$$\frac{1}{20} \frac{d\dot{m}_{fpf}}{dt} = P_9 - P_{fp} - B51 \frac{\dot{m}_{fpf}^2}{\rho_9} \tag{5.3}$$

is modifed as

$$\frac{1}{20}\frac{d\dot{m}_{fpf}}{dt} = P_9 - P_{fp} - .5B51\frac{\dot{m}^2_{fpf}}{\rho_9} \qquad (5.4)$$

where the parameter $B51$ is the lumped area of the fuel preburner flow area, $\dot{m}_{fpf}$ is the mass flow rate into the fuel preburner, P_{fp} is the fuel preburner pressure, and P_9 and ρ_9 are the pressure and fuel density in the preburner common fuel supply line respectively.

Example 3 : Failures in the turbopump bearings can be modeled as additional resistances or additional torques absorbed by the pump. For instance, bearing failure in the High Pressure Fuel Turbopump can be represented as changes in parameters $B41$ and $B44$ in the following equations

$$\tau_{op2} = 1.2B41\Omega^2_{o2}\Gamma(\phi_{op2}) \qquad (5.5)$$

$$\tau_{op3} = 1.2B44\Omega^2_{o2}\Gamma(\phi_{op3}) \qquad (5.6)$$

where τ_{op2} and τ_{op3} are the torques absorbed by the HPOTP and the boost pump respectively. The quantities $\Gamma(\phi_{op2})$ and $\Gamma(\phi_{op3})$ are nonlinear functions of other state variables [24]. The coefficients $1.2's$ in front of $B41$ and $B44$ represent a 20 per cent jumps in the torques absorbed by the pumps due to bearing failures.

Each of these modified equations (i.e. Eqs. 5.2, 5.4-5.6), representing a failure, changes the internal dynamics of the engine and produces a new nonlinear model. The problem of designing a FDI scheme for engine internal component failures becomes the problem of identifying a model from a set of models (each model has one or some parameters modified in the same fashion as Eqs.(5.2-5.6)) that best matches the model of the real failure. The identified model indicates the corresponding faulty subsystem. Using the MM technique is sufficient for this purpose as we demonstrate in the next sections.
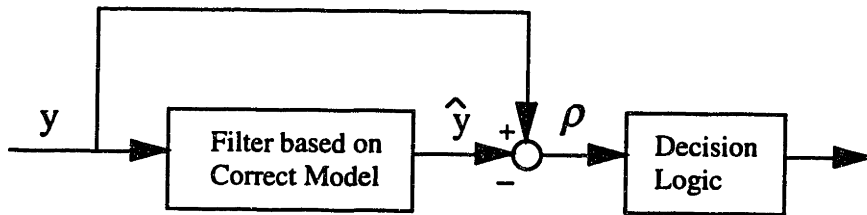
Figure 5-2: Residual-Based FDI Structure.
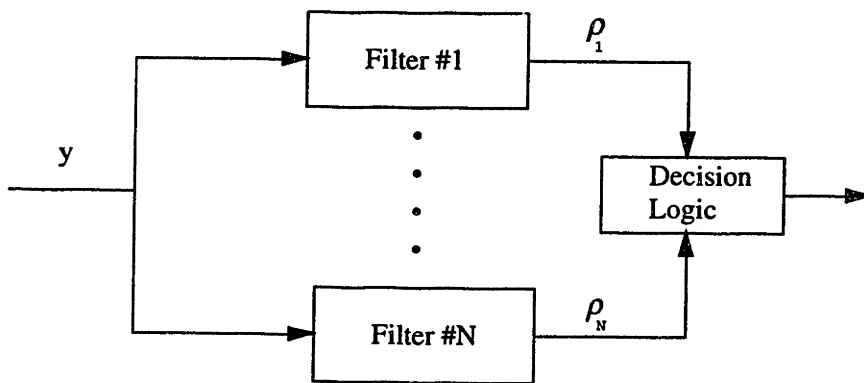
Figure 5-3: Multiple-Filter FDI Structure.

5.2 The Multiple Model Method

The FDI method presented in Chapter 4 for detection of failures in the sensors and valves has a residual-based structure as depicted in Figure 5-2, where only one Kalman filter based on the correct model is used in the implementation. The decision mechanism in this structure is based on the residual generated by the filter. For detection of failures in engine internal components, we introduce the Multiple Model(MM) method which has a multiple-filter structure as illustrated in Figure 5-3. In contrast with the single filter approach of the residual-based structure, the MM approach uses a set of N filters based on different hypothesized models to process the observed measurement, y. Each hypothesized model corresponds to an assumed model, i.e. Filter #1 assumes the correct model, Filter #2 assumes a model with changes in system parameters (i.e. such as that of parameter $B65$ in Eq.(5.2)) that correspond to a failure in a particular component, etc. The residual generated by each filter is small if the hypothesized model is close to the true model. Consequently the filter that generates smallest residual is one whose model best matches that of the true model. Hence, the decision logic is based on deciding which of these N hypothesized filters is performing best in term of keeping the residual small.

The formulation of the MM method presented here follows the presentation in [30]. A natural framework for formulating the MM problem is posing it in terms of a multiple hypothesis test, as done in Section 4.1. Specifically, by assuming that the true model belongs to the set of N models, we have the following multiple hypothesis test

$$H_i \quad : \quad x_i(k+1) \quad = \quad A_i x_i(k) + B_i u(k) + \Gamma_i w_i(k) \qquad (5.7)$$

$$y(k) \quad = \quad C_i x_i(k) + v_i(k) \qquad (5.8)$$

for $i = 1, 2, ..., N$. The matrices A_i, B_i, C_i, and Γ_i are known, $u(k)$ is the control input vector, $x_i(k)$ is the state vector, and each $x_i(k)$ can be of different

dimensions depending on how the hypothesized model is used to approximate the true model. The sequences $w_i(k)$ and $v_i(k)$ are independent, zero mean, Gaussian white noise with covariance $Q_i(k)$ and $R_i(k)$ respectively. In addition, they are independent of the initial state estimate $\hat{x}_i(0)$ that has covariance $P_i^+(0)$.

The corresponding Kalman filter equations for each model i are

$$\hat{x}_i^-(k+1) = A_i(k)\hat{x}_i^+(k) + B_i(k)u(k) \tag{5.9}$$

$$\hat{x}_i^+(k+1) = \hat{x}_i^-(k+1) + K_i(k+1)\varrho_i(k+1) \tag{5.10}$$

$$\varrho_i(k+1) = y(k+1) - C\hat{x}_i^-(k+1) \tag{5.11}$$

$$P_i^-(k+1) = A_i(k)P_i^+(k)A_i'(k) + \Gamma_i(k)Q_i(k)\Gamma_i'(k) \tag{5.12}$$

$$V_i(k) = C_i(k)P_i^-(k)C_i'(k) + R_i(k) \tag{5.13}$$

$$K_i(k) = P_i^-(k)C_i'(k)V_i(k)^{-1} \tag{5.14}$$

$$P_i^+(k) = P_i^-(k-1) - K_i(k)C_i(k)P_i^-(k-1) \tag{5.15}$$

where $\hat{x}_i^-(k)$ and $\hat{x}_i^+(k)$ are the predicted and update estimates of the state respectively, while $P_i^-(k)$ and $P_i^+(k)$ are their corresponding error covariance matrices. The matrix $K_i(k)$ is known as the Kalman gain and $V_i(k)$ is the covariance of the residual $\varrho_i(k)$.

Let $p_i(0)$ be the a priori probability that H_i is true, and likewise $p_i(k)$ be the probability that model i is the true model up to time k given $J_k = [u(0), ..., u(k), y(1), ..., y(k)]$. Using Bayes' rule, we can express $p_i(k+1)$ as

$$p_i(k+1) = \frac{p(y(k+1)|H_i, J_k)p_i(k)}{\sum_{j=1}^{j=N} p(y(k+1)|H_j, J_k)p_j(k)} \tag{5.16}$$

But from Eq.(5.11), $y(k+1)$ has the statistics of a Gaussian random vector with mean $C_i(k)\hat{x}_i^-(k)$ and covariance $V_i(k)$ under H_i. As a result, the conditional

probability density $p(y(k+1)|H_i, J_k)$ for $i = 1, 2, ..., N$ can be computed as

$$p(y(k+1)|H_i, J_k) = \frac{\exp[-\frac{1}{2}\varrho_i'(k+1)V_i^{-1}(k+1)\varrho_i(k+1)]}{(2\pi)^{m/2}[detV_i(k+1)]^{\frac{1}{2}}} \qquad (5.17)$$

where m is the dimension of y.

In summary, the implementation of the MM algorithm consists of using the inputs $y(k)$ and $u(k)$ to compute the probabilities of Eq.(5.17) to determine which model from a set of N hypothesized models best matches that of the true model, assuming that the system has reached steady-state. A bank of N filters is used in this process, where each filter uses one of these models. The performance of the filters is reflected in the magnitude and whiteness property of the residual $\varrho_i(k)$, i.e., if the ith model is the true model, then $\varrho_i(k)$ will be unbiased and white; conversely, if the the assumed model is not correct, $\varrho_i(k)$ will not be white and unbiased as a result of using the wrong model for estimation.

As discussed in [30], one important issue in the implementation of the MM method is how the algorithm handles the case when the real system under consideration is nonlinear. This problem can be remedied by using the extended Kalman filter (EKF) in place of the KF for each hypothesized model. The EKF is designed based on the linearized model about different operating points. The performance of the MM algorithm in this case relies on how far apart the linearized models are, which is a signal to noise ratio problem. Specifically, the quantity $V_i(k)$ dictates how well each filter is tracking. Large value of $V_i(k)$ indicates that the quantity $\varrho_i'(k)V_i^{-1}(k)\varrho_i(k)$ is small and in effect flattens out the probability density $p(y(k+1)|H_i, I(k), u(k))$, which makes it harder to distinguish among the models.

Another important issue is that in addition to identifying the correct model, the MM algorithm also needs to detect and track the shifting of the hypothesized models from one to another.

In our implementation of the MM method, we made the simplification that

Subsystem	Parameters and Their Description
HPOTP subsystem	B58 (lumped cross-section area of the flow passage in the HPOTP) B60B61 (lumped cross-section area of the HPOTP injector) B55 (HPOTP isentropic efficiency)
HPFTP subsystem	B51 (lumped cross-section area of the flow passage in the HPFTP) B53B54 (lumped cross-section area of the HPFTP injector) B8 (HPFTP isentropic efficiency)
LPOTP subsystem	B32(lumped cross-section area of flow passage in the LPOTP)
LPFTP subsystem	B12(lumped cross-section area of flow passage in the LPFTP)
Main Injector	B65(lumped cross-section area of the main injector posts)

Table 5.1: Selected parameters in each subsystem and parameters' descriptions.

the system approaches steady-state rapidly after a failure occurs, and therefore there is no need to track transients arising from shifting models. This assumption is particularly plausible because the SSME is a highly damped system.

5.3   Results

In this section, we present some numerical results from the implementation of the MM method. As noted in Section 5.1, our goal is to isolate the failures according to the subsystem in which they originate. To demonstrate this idea, a few parameters are chosen in each subsystem as listed in Table 5.1, and the failures are modeled as jumps in the parameters as shown in Section 5.1. Each failure is represented by a new model that is identical with the nominal model except with the modified parameter. For simplicity, each failure is modeled as a jump with a certain percentage in one parameter at a time. Hence, from Table 5.1, the model set for the hypothesis test consists of nine different models, one for each parameter. As shown in Figure 5-3, each of these new models is used to contruct a Kalman filter, and the MM algorithm computes the probability of which Kalman filter is realized by the correct model.

Figure 5-4 plots the probability of each of these hypothesized models/parameters that has failed. In this example, the failure consists of cracks

in the LOX posts of the HPFTP injector, which is modeled as a decrease by 50 per cent in parameter B53B54. Parameter B53B54 represents the lumped cross-section area of the HPFTP injector. After the transient response of approximately one second has died out, the algorithm correctly indicates that the probability (of the model that represents a jump by 50 per cent in parameter B5354) tends to one. Similar results are obtained for failures that correspond to jumps in the other parameters as shown in Figures 5-5, 5-6, and 5-7.

We want to emphasize here that the MM algorithm presented thus far assumes that the hypothesized model set consists of the true failed model along with other models. This assumption generally cannot be satisfied since it requires that the hypothesized model set must include a model that precisely matches with that of the real failed model. For example, in Figure 5-4, the real failure is a 50 per cent jump in parameter B5354 so the model set must include a model with precisely 50 per cent jump in parameter B5354. This assumption is not practical and can be relaxed as shown in the next section where the issue of how to choose hypothesized models to include in the hypothesis test is discussed.

5.4 Robustness of the MM method

One problem commonly encountered in practice using the MM algorithm is that the set of N hypothesized models inevitably does not include the true model due to changes in system parameters, unmodeled dynamics, modeling errors from linearization, etc. The criteria for choosing which models to include in the model set is not clear. A result had been derived for linear systems in [3] that introduces a *distance* between the models and shows that the MM algorithm will converge to the ith model that is closest to the true model according to that distance. In other words, the probability density function $p_i(k)$ of this i^{th} model will tend to *one*. Since the true model is not generally known when the dynamics of the system changes (i.e. changes in system parameters), this
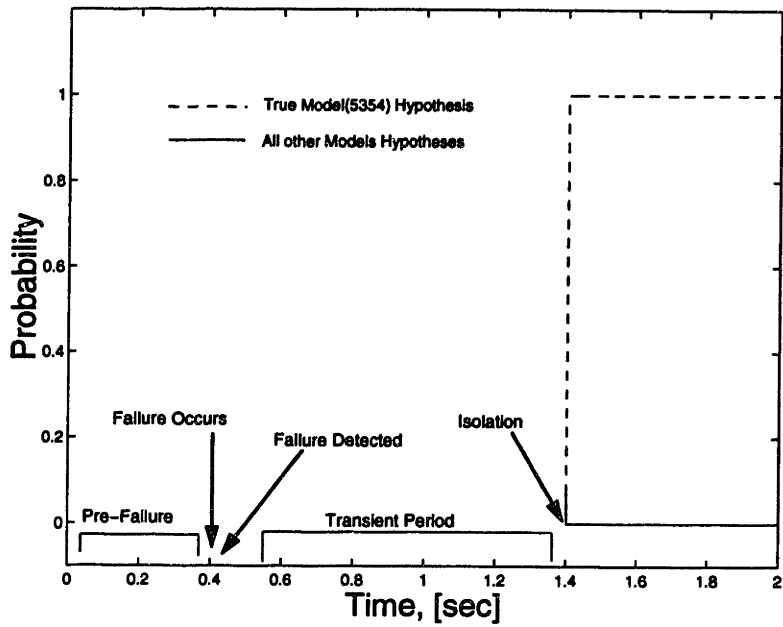
Figure 5-4: MM probability plot. Failure in Paramameters B51B53. Failure is isolated immediately after $t = 1.4$ seconds.
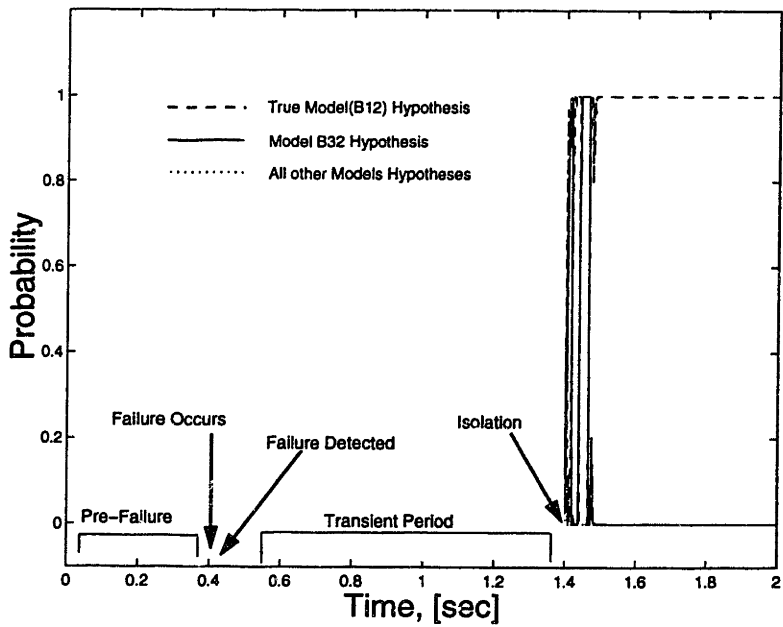
Figure 5-5: MM probability plot. Failure in Paramameter B12. Failure is isolated at $t = 1.47$ seconds.
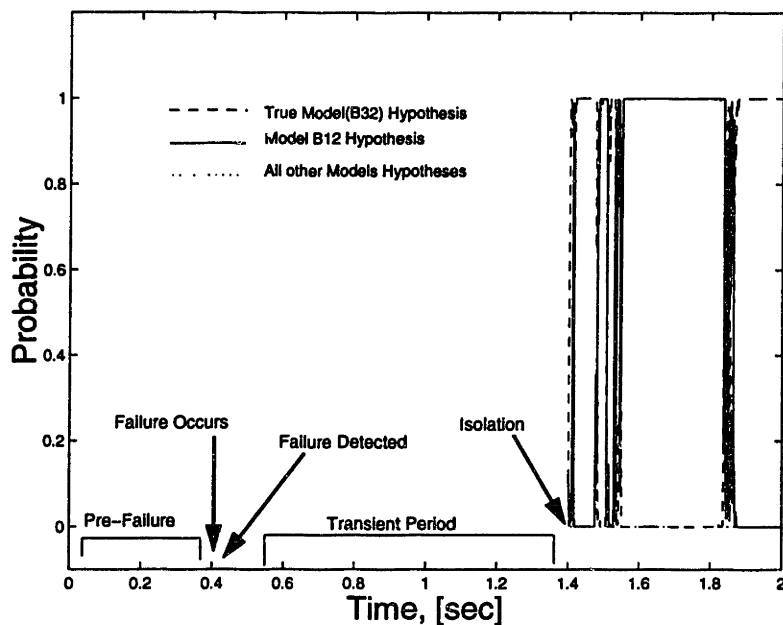
Figure 5-6: MM probability plot. Failure in Paramameter B32. Failure is isolated at $t = 1.86$ seconds.
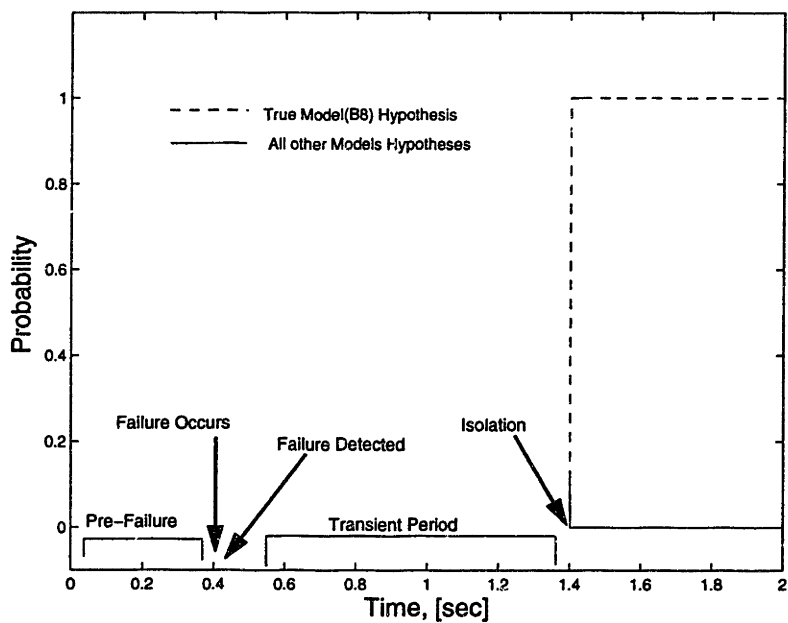
Figure 5-7: MM probability plot. Failure in Paramameter B8. Failure is isolated immediately after $t = 1.4$ seconds.

93

technique essentially provides a way to check the robustness of the MM method or a guide in selecting models far apart enough so that the MM algorithm will be able to distinguish between them.

The equations of the algorithm that are necessary for the calculation of the distance between the models are summarized below. A thorough treatment of the algorithm can be found in [3]. Suppose we have a model set of N models denoted by

$$\mathcal{M} \equiv \{M_j = (A_j, B_j, C_j, Q_j, R_j); j \in Z = (0, 1, ..., N)\} \qquad (5.18)$$

where the true model i is not included in the set Z. This algorithm states that the MM method will converge to model j whose distance, L_j^i, is closest to the true model i where $j \in Z$

$$|L_j^i| = \min\{|L_o^i|; o \in Z\} \qquad (5.19)$$

The distance L_j^i is computed by the following procedure. First we assume that the model M_k is detectable[1], all the eigenvalues of matrix A are inside the unit circle, and the dynamics equations as well as the filter equations have reached steady state. These conditions guarantee that the steady-state prediction error covariance matrix $P_j^-(k)$ and the ergodicity[2] of $\hat{y}^-(k)$ and $y_j(k)$ exist. The matrix $P_j^-(k)$ is generated by the filter corresponding to model M_j. Since we are at steady-state operation, the time step subscript k is dropped off from all matrices. Let

$$\Gamma_j^i \equiv E_i\{(y - \hat{y}_j^-)'(y - \hat{y}_j^-)\} \qquad (5.20)$$

be the prediction error covariance matrix where the expectation is taking with respect to model i while the measurement is from model j. The matrix Γ_j^i can

---

[1] A linear time invariant system is detectable if all unobservable modes are stable.

[2] A zero mean stationary Gaussian process $z(n)$ is ergodic if and only if $\lim_{n \to \infty} \frac{1}{n+1} \sum_{n=0}^{n} |E\{z(n)z'(n+k)\}|^2 = 0$.

be expressed as

$$\Gamma_j^i = C_j^i \Psi_j^i C_j^{\prime i} + R_i \qquad (5.21)$$

where Ψ_j^i is the steady-state solution of the Lyapunov equation

$$\Psi_j^i(k+1) = A_j^i \Psi_j^i(k) A_j^{\prime i} + G_j^i Q^i G_j^{\prime i} \qquad (5.22)$$

with arbitrary initial condition and

$$A_j^i = \begin{bmatrix} A_i & 0 \\ A_j K_j C_i & A_j(I - K_j C_j) \end{bmatrix} \quad ; \quad G_j^i = \begin{bmatrix} G_i & 0 \\ 0 & A_j K_j \end{bmatrix}$$

$$Q^i = \begin{bmatrix} Q_i & 0 \\ 0 & R_i \end{bmatrix} \quad ; \quad C_j^i = \begin{bmatrix} C_i & -C_j \end{bmatrix}$$

Then the distance between models j and i is

$$L_j^i = \log |P_j^-| + \mathrm{trace}(P_j^{-1}\Gamma_j^i) \qquad (5.23)$$

The calculation of the distance L_j^i consists of computing the prediction error covariance matrix Γ_j^i using Eqs.(5.21-5.22) and the prediction error covariance matrix P_j^- using the Kalman filter Eqs.(5.9-5.15).

To demonstrate the convergence of the MM algorithm to the model that best matches the real failed model, a simulation is presented in Figure 5-8 where the real failed model is a jump of 36 per cent in paramater $B8$. In this example, the model set Z consists of a model that represents a jump of 15 per cent in parameter $B8$ and other models that represent jumps of arbitrary percentage in other parameters. As shown in Figure 5-8, the MM algorithm correctly indicates that the model closest to the real model is the model that represents a jump of 15 per cent in parameter $B8$. This conclusion is confirmed by the results obtained from computing the distance between the hypothesized models and
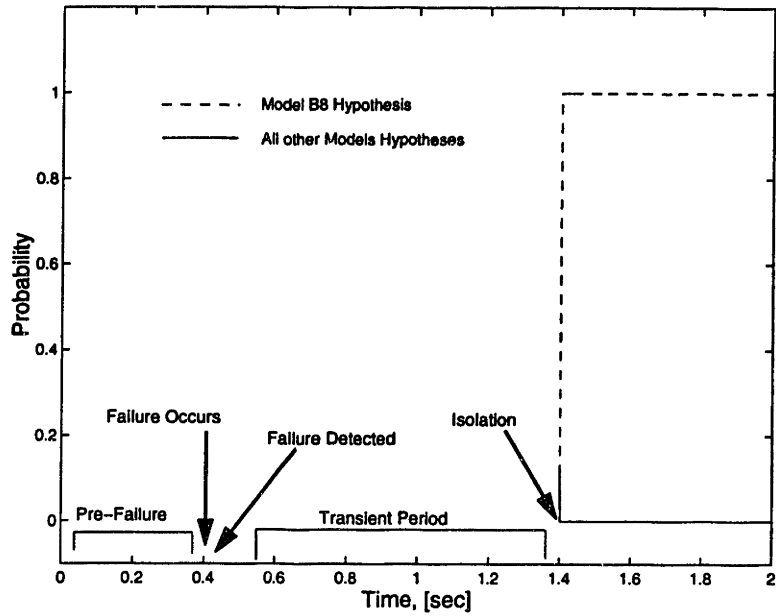
Figure 5-8: MM probability plot. Failure in Paramameter B8. The true failed model is not included in the model set. Failure is isolated immediately after t = 1.4 seconds.

the real model. Figure 5-9 shows the distance between the failed model and the hypothesized models where the $y - axis$ is the distance and the $x - axis$ is the hypothesized models. The shortest distance is the one between the real failed-model and the model with a 15 per cent jump in parameter $B8$. Similar results are obtained for failures in other models.

Thus, given the failures of interest, the distances between the models can be computed beforehand. The selected models are ones whose distances are far apart so that they can be isolated.
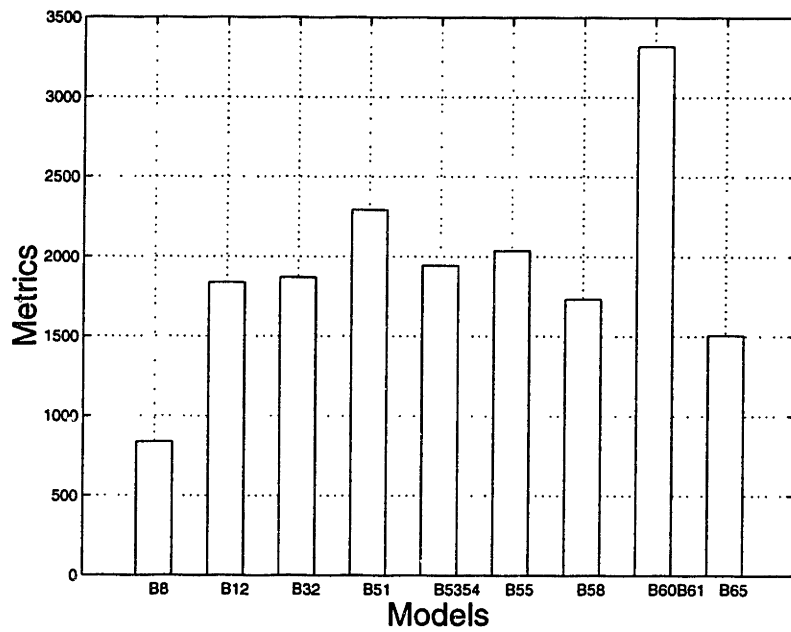
Figure 5-9: Distances (Metrics) between the real failed model and the hypothesized models.

Chapter 6

Conclusion

In this thesis, a Vehicle Health Monitoring (VHM) system is designed to detect and isolate failures in the engines of Reusable Launch Vehicles. In particular, our focus has been on the Space Shuttle Main Engine. This VHM system is an improvement over the current safety monitoring system that relies on red-line methods to detect failures. The algorithm of the VHM system is model based, combining information from a thermodynamic model of the engine together with sensor measurements to predict future sensor outputs. The residual, or the difference between the predicted and actual measurement, is used by several statistical tests to detect the presence of a failure and to identify failures in sensors and valves as well as internal components, such as turbopumps, main injectors, etc. The VHM system also supplies estimates of all of the state variables that can be used as a health record for the engine. This health record would help to reduce the cost of maintenance and routine checkup. In addition, at the design stage, the VHM system helps to determine where to best place sensors in order to obtain the necessary information to investigate the components in question.

6.1 Summary of Thesis

Since our FDI algorthm is model based, we first studied in Chapter 2 the description of the engine as well as the modeling of the engine dynamics. The SSME is an internal combustion cycle, and its dynamics can be categorized into seven types of thermodynamic equations. Since the SSME is a nonlinear model, linearization of the model was necessary to make the failure detection and isolation problem more tractable.

In Chapter 3, we proceeded to pose the problem of detecting and isolating engine failures as a hypothesis test. It was found that the Generalized Likelihood Ratio Test is difficult to apply due to the large number of potential failures and the nonlinearity of the engine model. These difficulties were circumvented by decomposing the problem into six steps: (1) linearize the nonlinear model, (2) design the Kalman filter based on the linearized model to generate a white innovation/residual sequence, (3) use the threshold test to process the residual and detect the presence of failures, (4) perform a whiteness test to distinguish sensor/valve failures and internal component failures, (5) employ the Generalized Likelihood Ratio Test (GLRT) to detect and isolate sensor/valve failures, and (6) employ the Multiple Model (MM) method to detect and isolate internal components. A description of steps (1)-(4) is also given in Chapter 3.

The linearization of the nonlinear model and designing the KF were straightforward. The threshold test, employing the T^2 statistic that is a function of the residual, decides whether a failure has occured by comparing the T^2 statistic against a threshold. Once a failure is detected, the whiteness test examines the autocorrelation function of the residual to check the whiteness of the residual. If the residual is white, then it is assumed that the failure is in either a sensor or a valve. If the residual is colored, then it is assumed that the failure is in an internal component.

Step (5) is described in Chapter 4. Specifically, the GLRT was implemented to accomodate failures that occur in sensors or valves. The GLRT uses a statis-

tical test to process the residual and identify the faulty component. Simulation results showed that failures in fifteen sensors and five valves, modeled as step failures, were successfully isolated using the GLRT algorithm. Besides step failures, the GLRT can be applied for other failure modes such as ramp and impulse. A geometrical interpretation of the GLRT was also presented to show that failures in the sensors and valves are distinguishable. This geometrical interpretation could be useful in the study of the issue of sensor placement to maximize the distinguishability of failures in the sensors and valves.

Finally, Chapter 5 deals with failures occurring in the engine internal components. Our goal was to determine in which subsystem (HPOTP, HPFTP, LPOTP, LPFTP, Main Injector, and Nozzle and pipe/ducts) the failure originated. This is a more realistic goal than tracking the precise origin of the failures. The internal component failures considered are those that occurred in the past, as reported in [8]. These failures, altering the geometry of the engine, were modeled as changes in the parameters of the nonlinear model since the geometry of the engine is represented by these parameters. Consequently, changes in parameters produce new models. A bank of Kalman filters were constructed based on these models. The Multiple Model (MM) method was used to process the residual of each of these filters and decide which filter corresponds to the correct model. Simulation results were presented for a few representative cases in which the failed parameters were identified correctly. When the true failed-model is not included in the model set (or not included in the bank of filters), we showed that the MM method converges to the model that is closest to the true failed model based on an information theoretic distance criterion.

6.2   Recommendations for Future Work

One immediate step that should be considered is testing the VHM algorithms with the engine real data. This step can be carried out in the following sequences.

1. Fine tune the nonlinear model with real data to ensure that the mathematical model is a reliable representation of the actual model. Various system identification techniques can be used in this process. For instance, a neural network can be implemented to identify the discrepancies (in term of modeling errors or unmodeled dynamics, etc.) between the actual model and the mathematical model.

2. Estimate the level of noise in the system and in the sensors based on the real data and the quality of the on-board sensors.

3. Test the VHM algorithms with real data to assess the performance of the KF, and set the thresholds appropriately based on the given information.

4. Consider other unmodeled dynamics in the system that the GLRT and the MM method may be sensitive to.

5. Find out the real set of on board sensors and their redundancy, i.e. which sensors and measurement instruments are triplicated, and modify the GLRT and the MM method accordingly.

6. The choice of parameters used in the MM method for internal component failures should be further analyzed meticulously to ensure that all failures of interested are modeled.

7. Test the VHM system with data containing the actual failures and reconfigure/adjust the designed algorithms.

With the algorithm working properly, the next step is to integrate the developed FDI algorithm into a larger VHM system. This integration requires fusing the VHM system with the current safety monitoring systems and perhaps with other health monitoring systems based on the techniques mentioned in Chapter 1.

Two other problems of equal importance that we have not discussed are the problem of reconfiguring the system to compensate the identified failure and

the problem of integrating the FDI algorithm into the closed-loop control of the engine. The former problem typically involves compensating for damages or loss of hardware functionality. For example, when a valve malfunctions or is stuck at one position, then one has to somehow use the other four valves to obtain the desired response. Another example is in the case when a sensor malfunctions, the filter has to be reconfigured to take into account the loss of this sensor while retaining the observability condition of the filter.

To integrate the FDI algorithm into the closed-loop control, one needs to design an interface (between the VHM system and the controller) that takes into account how the different features of the FDI algorithms and the control laws affect one another. As an example, the control laws are often designed to be insensitive to modeling error, which is a plus for the FDI algorithms, but the control laws may also be robust to failures that the FDI algorithm wants to detect and isolate. Another issue that should be considered is of augmenting the control states with the filter states to enhance the performance of both the filter and the controller.

Bibliography

[1] Hybrid Systems Analysis Unit and System Dynamics Unit, "Engine Balance and Dynamics Model," Rockwell International Coporation, Rocketdyne Division, report number RL00001, January 1992.

[2] Anderson, T. W., *An Introduction to Multivariate Statistical Analysis*, pp. 156-194, Second Edition, John Wiley and Sons, 1984.

[3] Baram, Y., *Information, Consistent Estimation and Dynamic System Identification*, Ph. D. Thesis, MIT, Nov. 1976.

[4] Basseville, M., Nikiforov, I., *Detection of Abrupt Changes, Theory and Application*, Prentice Hall, 1993.

[5] Benzing, D. A., Hopkins, R. C., Whitaker, K. W., *Experimental Verification of Neural Network-Based SSME Anomaly Detection*, AIAA/ASME/SAE/ASEE 33rd Joint Propulsion Conference and Exhibits, July 6-9, 1997, Seattle, Washington.

[6] Bickford, R. L., Bickmore, T. W., Meyer C. M., Zakrajsek J. F., *Real-Time Flight Data Validation for Rocket Engines*, AIAA/ASME/SAE/ASEE 32nd Joint Propulsion Conference, Lake Buena Vista, Florida, July 1-3, 1996.

[7] Cikanek, H. A., *Space Shuttle Main Engine Failure Detection*, IEEE Control Systems Magazine, Vol. 6, No. 3, 1986, pp. 601-611.

[8] Cikanek, H. A., *Characteristics of Space Shuttle Main Engine Software*, AIAA/SAE/ASME/ASEE 23rd Joint Propulsion Conference, San Diego, California.

[9] Duyar, A., Guo, T.-H., Merrill, W., Musgrave, J., *Implementation of a Model Based Fault and Diagnosis for the Space Shuttle Main Engine*, Prepared for the Third Annual Conference on Health Monitoring for Space Propulsion Systems, Cincinnati, Ohio, November 13-14, 1991.

[10] Duyar, A., Merrill, *Fault Diagnosis for the Space Shuttle Main Engine*, J. of Guidance, Control, and Dynamics, Vol. 15, No.2, March-April, 1992, pp. 384-389.

[11] Gelb A. (editor), *Applied Optimal Estimation*, MIT Press, 1974.

[12] Jackson, E., *Spacecraft Propulsion, Space Shuttle Main Engine Overview*, von Karman Institute for Fluid Dynamics, Lecture Series 1993-01, Rockwell International Corporation, Canoga Park, California, January 25-29, 1993.

[13] Jenkins, G. M., Watts, D. G., *Spectral Analysis and Its Applications*, Holden Day, San Francisco, 1968, pp. 187-189.

[14] Jianjun, W., Yulin, Z., and Qizhi, C., *A Real-time Verification System on Fault Diagnosis Methods for Liquid Propellant Rocket Engine*, 32nd AIAA/ASME/SAE/ASEE Joint Propulsion Conference, Lake Buena Vista, Florida, July 1996.

[15] Kumar, P. R., *Stochastic systems: estimation, identification, and adaptive control*, Prentice-Hall information and system sciences series, 1986.

[16] Lo, F., Wu, K., Whitehead, A. *Identification and Interpretation of Patterns in Rocket Engine Data*, Final Report Under NASA Marshall Space Flight Center NAG 8-166, Marshall Flight Center, Al 35812 July 1993.

[17] Mehra, R. K., Peschon, J., *An Innovation Approach to Fault Detection and Diagnosis in Dynamic System*, Automatica, Vol. 7, 1971, pp. 637-640.

[18] Draper, S., Mangoubi, R., Martinez-Sanchez, M., and Adams, N. *Vehicle Health Monitoring for the Space Shuttle Main Engine*, Charles Stark Draper Laboratory Report CSDL-R-2774 February 1997.

[19] Mangoubi, R., *Robust Estimation and Failure Detection: A Concise Treatment*, Springer Verlag Publishing, 1998.

[20] Massoumnia, M., *A geometric approach to failure detection and identification in linear systems*, Ph.D. Thesis, Department of Aeronautics and Astronautics, MIT, February 1986.

[21] Meyer, C., Maul, W., Dhawan, A., *SSME Parameter Estimation and Model Validity Using Radial Basis Function Networks*, Advanced Earth to Orbit Propulsion Technology Conference, May 1994.

[22] Nemeth, E., *Health Management System for Rocket Engines*, NASA Contractor Report 185223, June 1990.

[23] Panossian, H. V., Kemp, V. R., *Real-Time Fault Detection of the Space Shuttle Main Engine*, Proceedings of the 30th Conference on Decision and Control, December 1991, Brighton, England.

[24] Lozano, P., *Dynamic Models for Liquid Rocket Engines with Health Monitoring Application* M.S. Thesis, Department of Aeronautics and Astronautics, MIT, May 1998.

[25] Sutton, G., *Rocket Propulsion Elements, An Introduction to the Engineering of Rockets*, John Wiley and Sons, Inc., 1992, pp. 203.

[26] Tejwani, G. D., Bircher, F.E., Van Dyke, B.D., Thurman, C. C., *SSME Health Monitoring at SSC with Exhaust Plume Emission Spectroscopy*,

presented at the 33rd AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit, July 6-9, 1997, Seattle, WA.

[27] Van Trees, H. L., *Detection, Estimation, and Modulation Theory*, Part 1 and 2, John Wiley and Sons, Inc., 1968.

[28] Walker, B., and Huang, K.Y., *Fault Detection Performance of a FDI Strategy Based on Nonlinear Parameter Estimation for the Space Shuttle Main Engine*, presented at the 31st CDC, Tuscon, Arizona.

[29] Willsky, A., *6.433 Supplementary Notes: Recursive Estimation*, MIT, Cambridge, MA, Spring 1994.

[30] Willsky, A., *Detection of Abrupt Changes in Dynamic Systems*, Report LIDS-P-1351, Laboratory for Information and Decision Systems, MIT, Cambridge, MA, January 1984.

Appendix A

SSME Matrices and Equations

A.1 Linearized System Matrices and Noise Covariances

The results presented throughout this thesis are generated with the following system matrics and noise covariances. All the system matrices (A,B,C), in continuous time, are obtained by linearing the nonlinear model at 100 power level (or about the state xo and control uo) .

xo (37 x 1) =

  Rows 1 through 6

   2.8566e+03   9.6724e+01   6.9787e+03   6.6505e+01   2.3267e+01   5.0475e+03

  Rows 7 through 12

   3.2300e+03   8.0237e+02   3.0060e+03   3.5966e+03   6.5481e-01   6.5480e-01

  Rows 13 through 18

   3.2770e+03   7.8340e+01   4.9769e+03   3.4937e+01   1.0420e+03   2.8405e+01

  Rows 19 through 24

   8.8900e-04   1.2149e+03   4.6599e+02   1.0373e+03   5.2184e+01   1.0240e-03

Rows 25 through 30

  1.2603e+03   4.6577e+02   5.2184e+01   9.5575e-01   9.6877e-01   9.1550e-01

Rows 31 through 37

  5.4997e+03   6.1092e+01   1.7602e+02   9.9942e+01   5.2753e+02   1.6346e+03   8.9909e+02

uo (5 x 1) = [7.8231e-01   6.4107e-01   1.0000e+00   1.0000e+00   9.9999e-01]

A (37x37) =

Columns 1 through 6

-5.5861e+01	9.4944e+00	0	0	5.4523e+01	3.7486e+01
7.3801e+02	-2.3588e+03	-1.0000e+02	0	0	0
-1.7690e+01	3.8175e+04	0	-3.8120e+04	-3.8120e+04	0
0	0	2.0000e+00	-1.2039e+02	0	0
0	0	1.0000e+00	0	-1.6598e+02	-1.0000e+00
3.4983e+01	0	0	0	2.0339e+04	-1.9852e+02
-1.0516e+01	0	0	-4.3169e+02	2.2962e+02	5.9675e+01
1.1966e+02	-1.5728e+02	0	0	0	0
0	-4.5338e+03	0	-6.2470e+03	-8.2108e+03	0
0	0	0	-1.5079e+02	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	3.1699e+04	0	0
0	0	0	0	0	-1.0000e+01
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	1.0932e+01	0	7.5386e+00	9.9084e+00	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	1.1272e+01	0	7.7736e+00	1.0217e+01	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0

```
     0            0            0            0            0  -1.0000e-01
     0            0            0            0            0            0
     0            0            0            0            0            0
     0            0            0            0            0            0
5.7438e+0,   -7.0019e+00        0            0            0            0
     0   -1.5000e+02        0            0            0            0
     0   -2.9829e+00        0            0            0            0
     0            0            0            0            0            0
     0            0            0            0            0            0


Columns 7 through 12


-4.3475e+01   1.8589e+01        0            0            0            0
     0   -6.2912e+02        0            0            0            0
     0    1.3987e+01        0            0            0            0
     0            0            0            0            0            0
     0            0            0            0            0            0
1.2997e+02        0            0            0            0            0
-1.7227e+03        0    1.4347e+03   8.5797e+00        0            0
     0   -2.2610e+02  -2.5000e+01        0            0            0
3.8167e+03   3.4742e+03  -4.8922e+03        0            0            0
-3.4068e+01        0            0   -1.9390e+01        0            0
     0            0            0    3.5795e-01  -8.6087e+02        0
-1.0000e-01        0            0            0            0  -1.4352e+01
     0            0            0            0    2.8370e+05  -2.8370e+05
     0            0            0            0            0            0
5.4266e+02        0            0   -4.5337e+01        0            0
     0            0            0            0            0            0
6.2267e+00        0            0            0            0            0
     0            0            0    1.7898e+01        0            0
9.8310e-06        0            0            0            0            0
-4.6057e+00   1.2678e+00   7.5447e+00        0            0            0
     0            0            0            0            0            0
     0            0            0            0            0            0
     0            0            0    5.3693e+01        0            0
     0            0            0            0            0            0
-4.7492e+00   1.3073e+00   7.7798e+00        0            0            0
     0            0            0            0            0            0
     0            0            0            0            0            0
     0            0            0    3.5795e-01        0            0
     0            0            0    3.5795e-01        0            0
     0            0   -1.0000e-01   3.5795e-01        0            0
     0            0            0            0            0            0
     0            0            0    7.1590e-01        0            0
     0   -7.0019e+00        0            0            0            0
```

```
        0   -1.5000e+02              0              0              0              0
        0   -2.9829e+00              0              0              0              0
-1.3203e+01             0              0              0              0              0
        0              0              0              0              0              0
```

Columns 13 through 18

```
        0              0              0   -3.6315e+01              0              0
        0              0              0              0              0              0
        0              0              0              0              0              0
        0              0   -2.0000e+00              0              0              0
        0              0              0              0              0              0
        0              0              0    3.1147e+03              0              0
        0    3.6648e+02    1.5458e+02   -1.5290e+02    1.3388e+02              0
        0              0              0              0              0              0
        0    5.3033e+03              0    5.4682e+03   -5.0599e+01              0
        0    1.2801e+02    2.8992e+01              0              0   -1.5691e+02
-1.0000e-01             0              0              0              0   -8.1356e-01
 1.0000e-01             0              0              0              0              0
        0              0              0              0              0              0
        0   -2.6704e+02   -2.0000e+01              0              0              0
        0    2.6725e+03   -8.1684e+02              0              0              0
        0              0              0   -2.5910e+02              0              0
        0              0              0              0   -1.1905e+02    3.1502e+02
        0              0              0              0   -1.9211e+01   -5.0914e+02
        0              0              0              0   -3.7773e-05    1.0000e-03
        0   -6.3998e+00              0   -6.5988e+00    1.4081e+00   -3.5104e+02
        0              0              0              0    3.1913e+00   -1.9447e-03
        0              0              0              0              0              0
        0              0              0              0              0   -1.4522e+02
        0              0              0              0              0              0
        0   -6.5993e+00              0   -6.8044e+00    6.2964e-02              0
        0              0              0              0              0              0
        0              0              0              0              0              0
        0              0   -1.0000e-01              0              0   -9.6812e-01
        0              0              0              0              0   -9.6812e-01
        0              0              0              0              0   -9.6812e-01
        0   -1.0000e+04              0   -1.0000e+04              0              0
        0              0              0              0              0   -1.9362e+00
        0              0              0              0              0              0
        0              0              0              0              0              0
        0              0              0              0              0              0
        0              0              0              0    4.7066e+01   -4.1170e+01
        0              0              0              0              0              0
```

Columns 19 through 24

0	0	0	1.4232e+00	0	5.5704e+04
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	3.2299e+01	0	1.2642e+06
2.4070e+08	0	0	1.6006e+00	0	6.2647e+04
0	0	0	0	0	0
-1.9591e+06	0	0	-7.8417e+01	0	-3.0693e+06
0	0	0	1.2296e+00	-1.5691e+02	4.8127e+04
0	0	0	0	-8.1356e-01	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	7.0641e+06
0	0	0	1.0867e+02	0	4.2533e+06
0	0	0	0	0	3.0566e+06
-5.7367e+07	6.0385e+01	1.2078e+02	0	0	0
-2.5193e+07	0	0	0	-4.8406e+01	0
-7.9967e+01	0	0	0	0	0
5.4518e+04	-1.9022e+01	0	9.4630e-02	0	3.7039e+03
1.2356e+05	0	-8.8581e+00	0	0	0
0	0	0	-7.4219e+01	1.4355e+02	-6.4537e+06
0	0	0	-5.5134e+01	-3.6940e+02	-8.8210e+07
0	0	0	0	5.0000e-04	0
2.4378e+03	0	0	1.1141e+00	-1.9424e+02	4.3605e+04
0	0	0	5.7261e+00	-7.0033e-04	2.2412e+05
0	0	0	8.6698e+01	0	1.4066e+08
0	0	0	0	-9.6812e-01	0
0	0	0	0	-9.6812e-01	0
0	0	0	0	-9.6812e-01	0
0	0	0	0	0	0
0	0	0	0	-1.9362e+00	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
7.8980e+07	0	0	0	-4.1170e+01	0
0	0	0	0	0	0

Columns 25 through 30

0	0	1.4369e+01	0	0	0
0	0	0	0	0	0

0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	3.2611e+02	0	0	0
0	0	1.6160e+01	0	0	0
0	0	0	0	0	0
0	0	-7.9175e+02	0	0	0
0	0	1.2415e+01	-1.5691e+02	-1.5691e+02	-1.5691e+02
0	0	0	-8.1356e-01	-8.1356e-01	-8.1356e-01
0	0	0	0	0	0
0	0	0	0	0	0
0	0	-4.6279e+01	0	0	0
0	0	1.0972e+03	0	0	0
0	0	-2.0024e+01	0	0	0
0	0	0	0	0	0
0	0	0	-4.8406e+01	-4.8406e+01	-4.8406e+01
0	0	0	0	0	0
0	0	9.5545e-01	0	0	0
0	0	0	0	0	0
4.6684e+01	9.3279e+01	-2.8566e+02	0	0	0
0	0	0	-1.4522e+02	-1.4522e+02	-1.4522e+02
0	0	-5.0000e-04	0	0	0
-1.8424e+01	0	9.8522e-01	0	0	0
0	-1.5932e+01	0	0	0	0
0	0	-7.6597e+01	0	0	0
0	0	0	-2.0605e+02	-9.6812e-01	-9.6812e-01
0	0	0	-9.6812e-01	-1.9958e+02	-9.6812e-01
0	0	0	-9.6812e-01	-9.6812e-01	-6.5150e+02
0	0	1.0000e+04	0	0	0
0	0	0	-1.9362e+00	-1.9362e+00	-1.9362e+00
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	-4.1170e+01	-4.1170e+01	-4.1170e+01
0	0	0	0	0	0

Columns 31 through 37

3.3569e-02	-1.2274e+01	1.8589e+01	0	0	0	0
0	0	-5.8349e+02	1.0000e+02	1.9543e+02	0	0
0	0	1.3987e+01	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0

112

7.6187e-01	-2.7856e+02	0	0	0	0	0
3.7754e-02	-1.3804e+01	0	0	0	0	0
0	0	-1.4587e+02	2.5000e+01	4.8856e+01	0	0
-1.8497e+00	6.7630e+02	0	0	0	0	0
2.9003e-02	-1.6752e+02	0	0	0	0	0
0	-8.1356e-01	0	0	0	2.8463e-02	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
2.0000e+01	3.9531e+01	0	0	0	0	0
2.5632e+00	-9.3719e+02	0	0	0	0	0
1.0000e+01	1.7105e+01	0	0	0	0	0
0	0	0	0	0	0	0
0	-4.8406e+01	0	0	0	1.4231e+00	0
0	0	0	0	0	0	0
2.2321e-03	-8.1613e-01	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	-1.4522e+02	0	0	0	4.2694e+00	0
0	0	0	0	0	0	0
2.3017e-03	-8.4157e-01	0	0	0	0	0
0	0	0	0	0	0	0
-2.1190e+01	0	0	0	0	0	0
0	-9.6812e-01	0	0	0	2.8463e-02	0
0	-9.6812e-01	0	0	0	2.8463e-02	0
0	-9.6812e-01	0	0	0	2.8463e-02	0
0	1.0000e+04	0	0	0	0	0
-2.0000e-01	-5.1058e+00	0	0	0	5.6926e-02	0
0	0	-5.7239e+01	-5.4599e-10	-7.2716e-02	0	0
0	0	0	0	0	0	1.5000e+02
0	0	9.0111e+01	0	-2.4955e+01	0	0
0	-4.1170e+01	0	0	0	-1.6013e+01	0
0	0	0	-2.1570e+01	0	0	-2.4158e-01

B (37 x 5) =

0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
2.2496e+04	0	0	0	0
0	1.1702e+04	0	0	0
0	0	0	0	0
0	0	0	0	0

0	0	1.6431e+04	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	1.0474e+03	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	3.1422e+03	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	2.0948e+01	0
0	0	0	2.0948e+01	0
0	0	0	2.0948e+01	0
0	0	0	0	0
0	0	0	4.1896e+01	1.3274e+02
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0

C (15 x 37) =

Columns 1 through 6

1.0000e+00	0	0	0	0	0
0	1.0000e+00	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0

114

0	0	0	0	0	0
0	0	0	0	0	0
4.2055e+00	-5.4214e+00	0	0	0	0
6.4453e+00	-1.8627e+01	0	0	0	0
0	0	0	2.2596e+01	0	0
8.3251e-03	0	0	0	5.9978e+01	-8.1071e-02
0	0	0	0	0	0

Columns 7 through 12

0	0	0	0	0	0
0	0	0	0	0	0
1.0000e+00	0	0	0	0	0
0	0	1.0000e+00	0	0	0
0	0	0	1.0000e+00	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	-5.4214e+00	0	0	0	0
0	-5.4214e+00	0	0	0	0
1.3727e-01	0	0	0	0	0
1.1846e-01	0	0	0	0	0
0	0	0	3.0444e+00	0	0

Columns 13 through 18

0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	1.0000e+00	0	0	0
0	0	0	0	1.0000e+00	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	-2.0004e+01	-9.4359e-02	0	0	0
0	0	0	-4.0224e+01	0	0
0	0	0	0	0	-1.9888e+00

Columns 19 through 24

19	20	21	22	23	24
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
1.0000e+00	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	1.4949e-01	0	3.8829e+03
0	0	0	1.4438e-01	0	3.7500e+03
0	0	0	0	-1.9888e+00	0

Columns 25 through 30

25	26	27	28	29	30
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	2.6264e+00	0	0	0
0	0	2.5365e+00	0	0	0
0	0	0	-1.9888e+00	-1.9888e+00	-1.9888e+00

Columns 31 through 37

31	32	33	34	35	36	37
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0

116

0	0	0	0	0	0	0
0	0	0	0	1.0000e+00	0	0
0	0	0	0	0	1.0000e+00	0
0	0	-5.0664e+00	1.0000e+00	1.7150e+09	0	0
0	0	-5.0664e+00	1.0000e+00	1.7150e+00	0	0
3.9242e-03	-1.8818e+00	0	0	0	0	0
3.7900e-03	-1.8174e+00	0	0	0	0	0
0	-1.9888e+00	0	0	0	2.5988e-01	0
0	0	-5.0664e+00	1.0000e+00	1.7150e+00	0	0
0	0	-5.0664e+00	1.0000e+00	1.7150e+00	0	0
3.9242e-03	-1.8818e+00	0	0	0	0	0
3.7900e-03	-1.8174e+00	0	0	0	0	0
0	-1.9888e+00	0	0	0	2.5988e-01	0

R (15 x 15) = diag[10 .5 15 15 15 20 5 5e-6 2 8 20 30 8 6 30]

Q (37 x 37) = diag[150 5 125 3.5 1.5 250 150 40 150 150 .03 .03 150 4 250 1.5 50 1.5
5e-5 60 20 50 2.5 5e-5 60 25 2.5 .05 .05 .05 250 3 8 5 25 90 45]

G = Q

A.2   RHS of the SSME Equations

This section lists the variables on the right hand side of Eqs.(2.10-2.46) and their relationship to the state variables and the control variables. Explicit forms of these equations can be found in [1] and [24].

$$\tau_{ot2} = f(\dot{m}_{opf}, \dot{m}_{opo}, \Omega_{o2}, P_{fi}, P_{op})$$

$$\tau_{op2} = f(\dot{m}_{mov}, \dot{m}_{op3}, \dot{m}_{ot1}, \Omega_{o2})$$

$$\tau_{op3} = f(\dot{m}_{op3}, \Omega_{o2})$$

$$P_{od3} = f(\dot{m}_{mov}, \dot{m}_{op3}, \dot{m}_{ot1}, \Omega_{o2}, \Omega_{o1}, P_{os})$$

$$\dot{m}_{op2c} = f(\dot{m}_{mcv}, \dot{m}_{op3}, \dot{m}_{ot1}, \Omega_{o2}, \Omega_{o1}, P_{os})$$

$$\dot{m}_{ot2} = f(\dot{m}_{opf}, \dot{m}_{opo}, \Omega_{o2}, P_{fi}, P_{op})$$

$$\dot{m}_{ft1} = f(P_{fi}, \rho_5, SU_5)$$

$$\dot{m}_{ft2} = f(\dot{m}_{fnbp}, \dot{m}_{fpf}, \dot{m}_{fpo}, \dot{m}_4, \Omega_{f2}, P_{fi}, P_{fp}, \rho_4, SU_4)$$

$$\dot{m}_{fi} = f(\dot{m}_{fnbp}, \dot{m}_{fpf}, \dot{m}_{fpo}, \dot{m}_{opf}, \dot{m}_{opo}, \dot{m}_4, P_c, P_{fi}, \rho_5, SU_5, \rho_4, SU_4)$$

117

$$P_{od2} = f(\dot{m}_{mov}, \dot{m}_{op3}, \dot{m}_{ot1}, \Omega_{o2}, \Omega_{o1}, P_{os})$$

$$\dot{m}_{cn} = f(\dot{m}_{fnbp}, \dot{m}_{fpf}, \dot{m}_{fpo}, \dot{m}_{mov}, \dot{m}_{opf}, \dot{m}_{opo}, \dot{m}_{op3}, ...$$
$$\dot{m}_4, P_c, P_{fi}, \rho_5, SU_5, \rho_4, SU_4)$$

$$\tau_{ft2} = f(\dot{m}_{fnbp}, \dot{m}_{fpf}, \dot{m}_{fpo}, \dot{m}_4, \Omega_{f2}, P_{fi}, P_{fp}, \rho_4, SU_4)$$

$$\tau_{fp2} = f(\dot{m}_{asifp}, \dot{m}_{asimc}, \dot{m}_{asiop}, \dot{m}_{fn}, \dot{m}_{fnbp}, \dot{m}_{mc}, \Omega_{f2})$$

$$P_{fd2} = f(\dot{m}_{asifp}, \dot{m}_{asimc}, \dot{m}_{asiop}, \dot{m}_{fn}, \dot{m}_{fnbp}, \dot{m}_{mc}, \Omega_{f2}, \Omega_{f1})$$

$$\dot{Q}_{15} = f(\dot{m}_{mc}, \rho_5, SU_5, T_{w15})$$

$$\dot{Q}_{25} = f(\dot{m}_{mc}, \rho_5, SU_5, T_{w25})$$

$$\dot{m}_5 = f(P_{fi}, \rho_5, SU_5)$$

$$H_5 = f(\rho_5, SU_5)$$

$$P_{mfvd} = f([A/\bar{A}]_{mfv}, \dot{m}_{asifp}, \dot{m}_{asimc}, \dot{m}_{asiop}, \dot{m}_{fn}, \dot{m}_{fnbp}, \dot{m}_{mc}, \Omega_{f2}, \Omega_{f1})$$

$$P_5 = f(\rho_5, SU_5)$$

$$\dot{Q}_{tc5} = f(\dot{m}_{cn}, \dot{m}_{fi}, \dot{m}_{mov}, \dot{m}_{op3}, T_{w25})$$

$$\dot{Q}_{14} = f(\dot{m}_{fn}, \rho_4, SU_4, T_{w14})$$

$$\dot{Q}_{24} = f(\dot{m}_{fn}, \rho_4, SU_4, T_{w14})$$

$$H_4 = f(\rho_4, SU_4)$$

$$P_4 = f(\rho_4, SU_4)$$

$$C_{\dot{m}_{fn}} = f(\rho_4, S\ddot{U}_4)$$

$$\dot{Q}_{tc4} = f(\dot{m}_{fnbp}, \dot{m}_{fpf}, \dot{m}_{fpo}, \dot{m}_{mov}, \dot{m}_{opf}, \dot{m}_{opo}, \dot{m}_{op3}, \dot{m}_4, ...$$
$$P_c, P_{fi}, \rho_5, SU_5, \rho_4, SU_4, T_{w14})$$

$$C_{\dot{m}_4} = f(\rho_4, SU_4)$$

$$P_{od1} = f(\dot{m}_{mov}, \dot{m}_{op3}, \Omega_{o1}, P_{os})$$

$$\tau_{ot1} = f(\dot{m}_{ot1}, \Omega_{o1})$$

$$\tau_{op1} = f(\dot{m}_{mov}, \dot{m}_{op3}, \Omega_{o1})$$

$$\tau_{ft1} = f(\Omega_{f1}, P_{fi}, \rho_5, SU_5)$$

$$\tau_{fp1} = f(\dot{m}_{asifp}, \dot{m}_{asimc}, \dot{m}_{asiop}, \dot{m}_{fn}, \dot{m}_{fnbp}, \dot{m}_{mc}, \Omega_{f1})$$

$$P_{ot} = constant$$

Appendix B

Linearization Procedure

Consider the nonlinear differential equation

$$\frac{dz}{dt} = f(z, s) \qquad \text{(B.1)}$$

where $z \equiv z(t)$ is the state of the system, $s \equiv s(t)$ is the control input, and $f(z, s)$ describes the dynamics. The functions z, s are continuous in the time variable t, while f is continuous in z and s. Let $s(t) = s_o(t)$ be a particular control history, producing a state trajectory $z(t) = z_o(t)$, i.e.,

$$\frac{dz_o}{dt} = f(z_o, s_o) \qquad \text{(B.2)}$$

Consider now a perturbation $s' = s_o + \delta s$ around s_o. By continuity, the resulting trajectory is $z' = z_o(t) + \delta z(t)$. Substitute z' for z into Eq.(B.1)

$$\begin{aligned}
\frac{dz}{dt} &= \frac{d}{dt}(z_o + \delta z) \\
&= f(z_o + \delta z_o, s + \delta s_o) \qquad \text{(B.3)}
\end{aligned}$$

and expand the right-hand side in a Taylor series expansion, we get

$$f(z_o + \delta z_o, s + \delta s_o) = f(z_o, s_o) + \frac{\partial f}{\partial z}(z_o, s_o)\delta z_o + \frac{\partial f}{\partial s}(z_o, s_o)\delta s_o + \text{h.o.t} \qquad \text{(B.4)}$$

119

Finally, substitute the above equation into Eq.(B.1) and substract Eq.(B.2) results in

$$\frac{d}{dt}\delta z = \frac{\partial f}{\partial z}(z_o, s_o)\delta z + \frac{\partial f}{\partial s}(z_o, s_o)\delta s + \text{h.o.t.} \qquad (B.5)$$

With

$$x(t) \equiv \delta z(t) \qquad (B.6)$$

$$u(t) \equiv \delta s(t) \qquad (B.7)$$

$$A_c(t) \equiv \frac{\partial f}{\partial z}(z_o, s_o) \qquad (B.8)$$

$$B_c(t) \equiv \frac{\partial f}{\partial s}(z_o, s_o) \qquad (B.9)$$

and the h.o.t replaced by white noise, we have the continuous-time, time-varying linear system

$$\frac{d}{dt}x(t) = A_c(t)x(t) + B_c(t)u(t) + \Gamma_c(t)w(t) \qquad (B.10)$$

where $w(t)$ is a white Gaussian process with unit intensity, and $\Gamma_c(t)$ is a design choice. The term $\Gamma_c(t)w(t)$ can represent the uncertainty, such as that emanating from the linearization. When discretized, the above system takes the form

$$x(k+1) = A(k)x(k) + B(k)u(k) + \Gamma(k)w(k) \qquad (B.11)$$

Here k is the time step, $x(k)$ the state at time k, $u(k)$ the control, and $w(k)$ is a white Gaussian noise with unit intensity.

If the control $s_o(t)$ and state $z_o(t)$ above are constant, i.e., $s_o(t) = s_o$ and $z_o(t) = z_o$, which is the case when the objective is to regulate around some state value, then the partial derivatives in Eqs.(B.8-B.9) are constants, so that $A(t) = A$, $B(t) = B$, etc. In that case, we would have $A(k) = A$, $B(k) = B$, and $\Gamma(k) = \Gamma$. In other words, the resulting linear system is time-invariant. This is the case when the SSME is held at constant power level. Finally, we rewrite

120

the observations equation (2.2).

$$y(k) = Cx(k) + v(k) \qquad \text{(B.12)}$$

where $v(k)$ the sensor noise, which is assumed to be white and Gaussian with a certain covariance.

Appendix C

A Derivation of the Kalman Filter

This appendix presents a derivation of the Kalman filter [29]. We first present some fundamental results from Baysian estimation theory in Section C.1 and then relate these results to the Kalman filter in Section C.2. Other ways of deriving the same result can be found in many signal processing texts such as [4, 11, 19].

C.1   Some Results from Estimation Theory

Estimation theory is an extremely rich field, a complete account of this subject can be found in [27]. Here we only present some fundamental results which will be essential tools for our derivation of the Kalman filter in the next section.

From Baysian estimation theory, let x and y be random variables where x is the quantity of interest and y is the measurement contains information about x. The notation $p_{y|x}(Y|X)$ is used to denote the conditional probability density function of y given x, and likewise Y is the value that y takes on give X. In particular, given a noisy measurement vector y of some function of x, the conditional density function $p_{y|x}(Y|X)$ and the density function of y, $p_y(Y)$ can both be deduced given the measurement model. Using Bayes rule, the posterior

density for x given the measurement y is then

$$p_{x|y}(X|Y) = \frac{p_{y|x}(Y|X)p_x(X)}{p_y(Y)} \tag{C.1}$$

In the Baysian approach, the goal is to choose an estimator $\hat{x}(y)$ using some error cost criteria and the information pertaining to x from the observed measurement y (i.e. the density $p_{y|x}(Y|X)$). In this thesis, we use a common cost criterion, mean-square error:

$$\hat{x}(y) = \arg\min_a \int_{-\infty}^{+\infty} (x-a)'(x-a)p_{x|y}(X|Y)dx \tag{C.2}$$

Likewise, this estimator is called the minimum mean-square error estimator since it minimizes the mean-square error. The estimate $\hat{x}(y)$ can be easily obtained by taking the derivative of Eq.(C.2) with respect to a and setting the result to zero as follows

$$
\begin{aligned}
-2a\int_{-\infty}^{+\infty}(x-a)p_{x|y}(X|Y)dx &= 0 \\
a\int_{-\infty}^{+\infty}p_{x|y}(X|Y)dx &= \int_{-\infty}^{+\infty}xp_{x|y}(X|Y)dx \\
a &= \hat{x}(y) = E[x|y]
\end{aligned} \tag{C.3}
$$

The obtained solution for a is indeed the global minimum since the cost function(Eq.(C.2)) is convex. Furthermore, this estimator is unbiased, and the estimation error covariance, Λ_e, is the expected value of the covariance of the posterior density

$$E[x-\hat{x}(y)] = E(x-E[x|y]) = E[x]-E[x] = 0 \tag{C.4}$$

$$\Lambda_e = E[(E[x|y]-x)(E[x|y]-x)'] = E[\Lambda_{x|y}(y)] \tag{C.5}$$

where $\Lambda_{x|y}(y)$ is the covariance of $p_{x|y}(X|Y)$. Another additional important property of this estimator is that its estimator error is uncorrelated with any function of the data $f(y)$ (i.e. $E[(x-\hat{x}(y))f'(y)] = 0$ or $E[xf'(y)] = $

123

$E[\hat{x}(y)f'(y)])$. This fact is known as the *orthogonality property*, and the result can be proven in a few steps

$$E[xf'(y)] = E[E[xf'(y)|y]] = E[E[x|y]f'(y)] = E[\hat{x}(y)f'(y)] \qquad (C.6)$$

The computation of this estimator, however, is not trivial since the estimator is generally a nonlinear function of the measurement y and the statistics of the density $p_{x|y}(X|Y)$ must be known. A situation when this estimator can be expressed in a simple form is the case when the vector x and y is jointly Guassian. Specifically, we have

$$\hat{x}(y) = m_x + \Lambda_{xy}\Lambda_y^{-1}(y - m_y) \qquad (C.7)$$

$$\Lambda_e = \Lambda_x - \Lambda_{xy}\Lambda_y^{-1}\Lambda_{xy}' \qquad (C.8)$$

where $m_x = E[x]$, $m_y = E[y]$, Λ_{xy} is cross-covariance of x and y, Λ_x and Λ_y is the covariance of x and y respectively. The estimator $\hat{x}(y)$ in this case is usually referred to as the linear least square estimator (LLSE). Eqs.(C.7-C.8) can be derived from [15] as follows. Let $\hat{z} = m_x + \Lambda_{xy}\Lambda_y^{-1}(y - m_y)$, and $\tilde{z} = x - \hat{z}$, . Then $E[\tilde{z}] = 0$ by direct substitution, and

$$
\begin{aligned}
E[\tilde{z}(y - m_y)] &= E[(x - m_x)(y - m_y)'] \\
&\quad -\Lambda_{xy}\Lambda_y^{-1}E[(y - m_y)(y - m_y)'] \\
&= \Lambda_{xy} - \Lambda_{xy}\Lambda_y^{-1}\Lambda_y \\
&= 0 \qquad (C.9)
\end{aligned}
$$

which implies that $\tilde{z}$ and y is uncorrelated or independent since they are Guassian, i.e. $E[\tilde{z}y] = 0$. Thus

$$
\begin{aligned}
\hat{x}(y) &= E[x|y] \\
&= E[\hat{z} + \tilde{z}|y]
\end{aligned}
$$

124

$$= E[\hat{z}|y] + E[\tilde{z}|y]$$

$$= \hat{z} \tag{C.10}$$

since $E[\tilde{z}|y] = E[\tilde{z}] = 0$. Therefore this proves Eq.(C.7). The proof for Eq.(C.8) is preceeded by noting that $\hat{z}$ and $\tilde{x} = x - \hat{x}(y)$ are independent because of the orthogonality property (C.6). This implies that $\Lambda_x = \Lambda_{\hat{x}} + \Lambda_{\tilde{x}}$. But $\Lambda_{\tilde{x}} = \Lambda_{x|y}$ by observing that the density $p_{\tilde{x}} \sim exp[-(x - m_{x|y})'\Lambda_{x|y}^{-1}(x - m_{x|y})]$ and the density $p_{x|y} \sim exp[-(x - m_{x|y})'\Lambda_{x|y}^{-1}(x - m_{x|y})]$ are identical. As a result,

$$\begin{aligned} \Lambda_{x|y} &= \Lambda_x - \Lambda_{\hat{x}} \\ &= \Lambda_x - \Lambda_{xy}\Lambda_y^{-1}\Lambda_{xy}' \end{aligned} \tag{C.11}$$

Using Eqs.(C.7-C.8), we can write down the LLSE equations for some simple expressions below which will be useful later in constructing the structure of the Kalman filter. Let

$$z = Fx + b + Gw \tag{C.12}$$

$$y = Cx + v \tag{C.13}$$

where F, G, C are known matrices, b is a deterministic vector, y is the observation vector, x and w and v are uncorrelated random vectors. Moreover, w and v has zero mean with covariances Q and R respectively. The vector z has the following prior statistics

$$m_z = Fm_x + b \tag{C.14}$$

$$\Lambda_z = F\Lambda_x F' + G\Lambda_w G' \tag{C.15}$$

Defined respectively $\hat{x}(y)$ and $\hat{z}(y)$ as the LLSE of x and z based on y, and $\Lambda_{x,l}$ and $\Lambda_{z,l}$ as the corresponding error covariances, applying Eqs.(C.7-C.8) and

straigth forward calculations yield

$$\hat{z}(y) = F\hat{x}(y) + b \tag{C.16}$$

$$\Lambda_{z,l} = F\Lambda_{x,l}F' + G\Lambda_w G \tag{C.17}$$

$$\hat{x}(y) = m_x + K(y - Cm_x) \tag{C.18}$$

$$K = \Lambda_x C'[C\Lambda_x C' + R]^{-1} \tag{C.19}$$

$$\Lambda_{x,l} = \Lambda_x - \Lambda_{xy}\Lambda_y^{-1}\Lambda'_{xy}$$

$$= \Lambda_x - \Lambda_x C'[C\Lambda_x C' + R]^{-1}C\Lambda_x \tag{C.20}$$

C.2   Kalman Filter Equations

We now present the derivation of Eqs.(3.21-3.27). The prediction Eqs.(3.21-3.22) are precisely Eqs.(C.16-C.17) since the model (3.15-3.16) is analogous to that of (C.12-C.13) and $w(k)$ is uncorrelated with $y(k)$, and $B(k)u(k)$ is a deterministic vector. The update equations are obtained in the same manner by noting the similarity among Eqs.(3.23-3.27) and Eqs.(C.18-C.20).

Thus the structure of the Kalman filter can be constructed immediately once the state space model is available. The only parameter to determine is the Kalman gain matrix which is computed together with the solution of the Riccati equation. In addition, one can see that the Kalman filter is essentially the solution to a sequence of static well-known linear least square estimation problems.

THESIS PROCESSING SLIP

FIXED FIELD: ill. _____ name _____

index _____ biblio _____

▶ COPIES: (Archives) Aero Dewey (Eng) Hum

Lindgren Music Rotch Science

TITLE VARIES: ▶ ☐ _____

_____

_____

NAME VARIES: ▶ ☐ _____

_____

IMPRINT: (COPYRIGHT) _____

▶ COLLATION: 126 p. _____

_____

▶ ADD. DEGREE: _____ ▶ DEPT.: _____

SUPERVISORS: _____

_____

_____

_____

_____

_____

_____

_____

NOTES:

cat'r: date:

▶ DEPT: _M.E._   page: ▶ 5/32

▶ YEAR: _1998_   ▶ DEGREE: _Ph.D. S.M._

▶ NAME: _HO, Nhut Tan_