Università degli Studi di Ferrara

DOTTORATO DI RICERCA IN
SCIENZE DELL'INGEGNERIA

CICLO XXI

COORDINATORE Prof. TRILLO STEFANO

DESIGN AND ANALYSIS OF LINEAR AND NONLINEAR

FILTERS FOR THE FDI OF AIRCRAFT MODEL SENSORS

Settore Scientifico Disciplinare ING-INF/04

Dottorando
Dott. BENINI MATTEO

Tutore
Prof. SIMANI SILVIO

Anni 2006/2008

Abstract

Increasing demands on reliability for safety critical systems such as aircraft or spacecraft require robust control and fault diagnosis capabilities as these systems are potentially subjected to unexpected anomalies and faults in actuators, input-output sensors, components, or subsystems. Consequently, fault diagnosis capabilities and requirements for aerospace applications have recently been receiving a great deal of attention in the research community.

A fault diagnosis system needs to detect and isolate the presence and location of the faults, on the basis also of the control system architectures. Development of appropriate techniques and solutions for these tasks are known as the fault detection and isolation (FDI) problem. Several procedures for sensor FDI applied to a nonlinear simulated model of a commercial aircraft, in the presence of wind gust disturbances and measurement errors, are presented in this thesis.

The main contributions of this work are related to the design and the optimisation of two FDI schemes based on a linear polynomial method (PM) and the nonlinear geometric approach (NLGA). In the NLGA framework, two further FDI techniques are developed; the first one relies on adaptive filters (NLGA–AF), whilst the second one exploits particle filters (NLGA–PF).

The suggested design approaches leads to dynamic filters, the so–called residual generators, that achieve both disturbance decoupling and robustness properties with respect to modelling errors and noise. Moreover, the obtained results highlight a good trade-off between solution complexity and achieved performances.

The FDI strategies are applied to the aircraft model in flight conditions characterised by tight–coupled longitudinal and lateral dynamics. The robustness and the reliability properties of the residual generators related to the considered FDI techniques are investigated and verified by simulating a general aircraft reference trajectory.

Extensive simulations exploiting the Monte–Carlo analysis tool are also used for assessing the overall performance capabilities of the developed FDI schemes in the presence of both measurement and modelling errors. Comparisons with other disturbance–decoupling methods for FDI based on neural networks (NN) and unknown input Kalman filter (UIKF) are finally reported.

Abstract (Italian)

Il bisogno crescente di affidabilità per sistemi critici dal punto di vista della sicurezza come quelli aeronautici e aerospaziali ha richiesto lo sviluppo di sistemi di controllo robusto e diagnosi dei guasti. Tali sistemi infatti sono potenzialmente soggetti a malfunzionamenti improvvisi, come guasti sugli attuatori, sui sensori di ingresso–uscita e sui componenti o sottosistemi. Di conseguenza, il problema della diagnosi dei guasti nell'ambito di applicazioni aeree e aerospaziali ha recentemente suscitato un notevole interesse nel settore della ricerca.

Un sistema automatico per la diagnosi dei guasti deve rivelare ed isolare la presenza dei guasti, anche sulla base della struttura del sistema di controllo. Lo sviluppo di tecniche e soluzioni adeguate per lo svolgimento di questi compiti rappresenta il problema fondamentale della diagnosi e dell'isolamento dei guasti (FDI). In questa tesi vengono presentate diverse procedure per la diagnosi dei sensori applicata a un modello simulato non lineare di un velivolo commerciale, in presenza di disturbi dovuti a raffiche di vento ed errori di misurazione.

I maggiori contributi di questo lavoro riguardano il progetto e l'ottimizzazione di due schemi di diagnosi basati su un metodo lineare polinomiale (PM) e su un approccio non lineare geometrico (NLGA). Nell'ambito del metodo NLGA, vengono successivamente sviluppate due ulteriori tecniche per la diagnosi dei guasti, la prima basata su filtri adattativi (NLGA–AF), mentre la seconda sfrutta i filtri particellari (NLGA–PF).

Le metodologie proposte hanno portato al progetto di filtri dinamici, i cosiddetti generatori di residuo, che permettono di ottenere sia il disaccoppiamento del disturbo, che interessanti proprietà di robustezza nei confronti degli errori di modellazione e del rumore. Inoltre, i risultati ottenuti evidenziano un buon compromesso tra la complessità della soluzione e le prestazioni ottenute.

Le strategie teoriche di diagnosi sono applicate al modello del velivolo in condizioni di volo caratterizzate da forte accoppiamento tra dinamiche longitudinale e laterale. Successivamente, le proprietà di robustezza e di affidabilità dei generatori di residuo così ottenuti sono state analizzate e verificate simulando una generica traiettoria di volo.

Numerose simulazioni sono state esaminate anche mediante lo strumento dell'analisi Monte–Carlo, necessario per valutare le prestazioni complessive delle tecniche di diagnosi proposte, in presenza di errori di misura e modellazione. Infine i metodi sviluppati sono confrontati con altre strategie di diagnosi con disaccoppiamento del disturbo basate sull'impiego di reti neurali (NN) e filtri di Kalman ad ingressi non noti (UIKF).

Acknowledgments

For their contribution to the realisation of this work I would like to thank

- Dr. Silvio Simani
- Dr. Marcello Bonfè
- Prof. Sergio Beghelli

from University of Ferrara and

- Dr. Paolo Castaldi
- Dr. Walter Geri

from University of Bologna.

Contents

1 Introduction **15**

2 Aircraft Simulation Model **19**
2.1 6 DoF Aircraft Model 19
2.1.1 Force Equations 20
2.1.2 Moment Equations 20
2.1.3 Euler Angles 21
2.1.4 True Air Speed and Aerodynamic Angles 23
2.1.5 Overall Model 24
2.2 Simulation Model Subsystems 26
2.2.1 Engine Model 26
2.2.2 Atmosphere Model 27
2.2.3 Servo–Actuators Model 28
2.2.4 Measurement Errors Description 29
2.2.5 NGC System 32
2.3 Aircraft FDI Model 33
2.3.1 PM FDI Model 33
2.3.2 NLGA FDI Model 34

3 Polynomial Method **37**
3.1 Residual Generation 37
3.1.1 Polynomial Basis Computation 38
3.1.2 Faults on the Input–Output Sensors 40
3.2 Residual Optimisation 40
3.2.1 Maximisation of the Steady–State Gain 42
3.2.2 Poles and Zeros Assignment 45
3.3 FDI on Input–Output Sensors 49
3.3.1 Bank for Input Sensors FDI 49
3.3.2 Bank for Output Sensors FDI 50
3.3.3 Fault Signature 52

4 Nonlinear Geometric Approach **55**
4.1 NLGA FDI Scheme . . . 55
4.1.1 Coordinate Transformation . . . 55
4.1.2 Residual Generators Design . . . 57
4.2 NLGA Robustness Improvements . . . 62
4.2.1 The $\bar{x}_{11}$–subsystem . . . 62
4.2.2 Filter Form and Observer Form . . . 63
4.2.3 Mixed $\mathcal{H}_-/\mathcal{H}_\infty$ Optimisation . . . 65
4.3 NLGA–AF FDI Scheme . . . 68
4.3.1 Adaptive Filtering Algorithm . . . 68
4.3.2 Adaptive Filters Design . . . 71
4.4 NLGA–PF FDI Scheme . . . 73
4.4.1 Basic Particle Filter Theory . . . 73
4.4.2 Throttle Particle Filter Design . . . 75

5 Simulation Results **77**
5.1 Threshold Test . . . 77
5.2 FDI Procedure for a Complete Trajectory . . . 78
5.3 Simulations and Performances Evaluation . . . 79
5.3.1 PM . . . 80
5.3.2 NLGA . . . 83
5.4 Comparative Studies and Robustness Evaluation . . . 88
5.5 Monte–Carlo Analysis . . . 91

6 Conclusion **95**

A Review of Model–Based FDI **99**
A.1 Basic Definitions . . . 99
A.2 Modelling of Faulty Systems . . . 101
A.3 Residual Generator General Structure . . . 104
A.4 Fault Detectability and Isolability . . . 106
A.5 Residual Generation Techniques . . . 107
A.6 Residual Evaluation . . . 108
A.7 Robustness Problem . . . 109

List of Figures

2.1 Aircraft axes and angles. 23
2.2 Overall architecture of the NGC system. 32

3.1 Graphical solution of Problem 1 when $q = 2$. 43
3.2 Automatic procedure to solve Problem 2. 48
3.3 Bank of filters for fault isolation on the input sensors. 50
3.4 Bank of filters for fault isolation on the output sensors. 51

5.1 Bank residuals for the 1–st input sensor fault isolation. 81
5.2 Bank residuals for the 9–th output sensor fault isolation. 82
5.3 NLGA elevator sensor FDI. 85
5.4 NLGA–AF elevator sensor FDI with fault size estimation. 85
5.5 NLGA and NLGA–PF residuals for throttle sensor FDI. 86
5.6 Comparison between NLGA–AF and fault estimator of Kaboré and Wang. 88
5.7 Aircraft complete trajectory example. 89

A.1 Structure of model–based FDI system. 100
A.2 Fault diagnosis in a closed–loop system. 101
A.3 The monitored system and fault topology. 102
A.4 Fault topology with actuator input signal measurement. 103
A.5 Residual generator general structure. 105
A.6 Residual generator using input–output description. 105

List of Tables

2.1 Wind gusts model parameters. 28
2.2 Transfer function poles of the servos. 29
2.3 Input sensor errors parameters. 29

3.1 Fault signatures. 52

5.1 PM minimal detectable step input sensor faults. 82
5.2 PM minimal detectable step output sensor faults. 83
5.3 NLGA minimal detectable step input sensor faults. 86
5.4 NLGA-AF minimal detectable step input sensor faults. 87
5.5 NLGA-PF minimal detectable step throttle sensor fault. 87
5.6 Performances for a complete aircraft trajectory. 90
5.7 PM Monte–Carlo analysis with $\nu = 4$ and $p = 10\%$. 92
5.8 NLGA Monte–Carlo analysis with $\nu = 12$ and $p = 10\%$. 92
5.9 NN Monte–Carlo analysis with $\nu = 5$. 93
5.10 UIKF Monte–Carlo analysis with $\nu = 9$. 93

Chapter 1

Introduction

The problem of Fault Detection and Isolation (FDI) in aircraft and aerospace systems has attracted considerable attention world–wide and been theoretically and experimentally investigated with different types of approaches, as can be seen from the general survey works (Gertler 1998, Chen and Patton 1999, Isermann 2005, Ding 2008). This development has been mainly stimulated by the trend in automation toward systems with increasing complexity and the growing demands for fault tolerance, cost efficiency, reliability, and safety which constitute fundamental design features in modern control systems.

Sensors are the most important components for flight control and aircraft safety and, as they work in a harsh environment, fault probabilities are high thus making these devices the least reliable components of the system. In order to improve the reliability of the system sensors hardware and software (analytical) redundancy schemes have been investigated over the last twenty years (Chen and Patton 1999, Isermann 2005).

For small aircraft systems, as considered in this work, multiple hardware redundancy is harder to achieve due to lack of operating space. Such schemes would also be costly and very complex to engineer and maintain. Analytical redundancy makes use of a mathematical model of the monitored process and is therefore often referred to as the model–based approach to FDI (Marcos *et al.* 2005, Amato *et al.* 2006). The model–based FDI is normally implemented as a computer software algorithm. The main problem of the model–based approach regards the real complex systems, where modelling uncertainty arises inevitably, because of process noise, parameter variations and modelling errors. The FDI of incipient faults represents a challenge to model–based FDI techniques due to inseparable mixture between fault effects and modelling uncertainty (Isermann 2005, Chen and Patton 1999).

A common and important approach in model–based techniques is known as the residual–based method. A number of researchers have developed residual–based methods for dynamic systems such as the parity space (Gertler 1998), state estimation (Basseville and Nikiforov 1993), Unknown Input Observer (UIO) and Kalman Filters (KF) (Chen and Patton 1999) and parameter identification (Basseville and Nikiforov 1993). Intelligent techniques (Korbicz *et al.* 2004) can be also exploited. Furthermore, the Massoumnia's geometric method (Massoumnia 1986) was successfully extended to nonlinear systems

(Hammouri *et al.* 1999, De Persis and Isidori 2001).

A crucial issue with any FDI scheme is its robustness properties. The robustness problem in FDI is defined as the maximisation of the detectability and isolability of faults together with the minimisation of the effects of uncertainty and disturbances on the FDI procedure (Chen and Patton 1999, Isermann 2005). However, many FDI techniques are developed for linear systems. Unfortunately, practical models in real world are mostly nonlinear. Therefore, a viable procedure for practical application of FDI techniques is really necessary. Moreover, robust FDI for the case of aircraft systems and applications is still an open problem for further research.

This work deals with the residual generator design for the FDI of input–output sensors of a general aviation aircraft, subject to wind gust disturbances and measurement noises. Two different FDI schemes are developed: the Polynomial Method (PM) and the NonLinear Geometric Approach (NLGA).

The developed PM scheme belongs to the parity space approach (Gertler 1998, Gertler and Singer 1990, Patton and Chen 1994) and it is based on an input–output polynomial description of the system under diagnosis. In particular, the use of input–output forms allows to easily obtain the analytical description for the disturbance decoupled residual generators. An appropriate choice of their parameters allows to maximise a suitable fault sensitivity function and to obtain desired transient properties in terms of a fault to residual reference transfer function. These dynamic filters, organised into bank structures, are able to achieve fault isolation properties.

The development of NLGA methodology is based on the works by De Persis and Isidori (De Persis and Isidori 2001). It was shown that the problem of the FDI for nonlinear systems is solvable if and only if there is an unobservability distribution that leads to a quotient subsystem which is unaffected by all faults but one. If such a distribution exists, an appropriate coordinate transformations in the state–space can be exploited for designing a residual generator only for the observable subsystem. The NLGA residual generators have been designed in order to be analytically decoupled from the vertical and lateral components of the wind. Moreover, a full analytical developed mixed $\mathcal{H}_-/\mathcal{H}_\infty$ optimisation is proposed, in order to design the NLGA residual generators so that a good trade–off between the fault sensitivity and the robustness with respect to measurements and model errors is achieved.

Two FDI techniques exploiting the NLGA coordinate transformations are also proposed: the NLGA–AF (Adaptive Filter) and the NLGA–PF (Particle Filter). The first one provides both FDI and the estimation of the fault size; it relies on the development of adaptive filters, instead of residual generators, for the observable subsystem obtained by the NLGA coordinate transformation. The second one, exploits particle filters to solve the FDI problem for the nonlinear stochastic model of the system under diagnosis, which is derived by following a NLGA strategy.

A very accurate flight simulator (simulation model) of the PIPER PA–30 aircraft, implemented in the Matlab/Simulink® environment, has been used to evaluate the effectiveness of the proposed method. The simulation model is based on the classical nonlinear

6 Degrees of Freedom (6 DoF) rigid body formulation (Stevens and Lewis 2003), whose motion occurs as a consequence of applied forces and moments (aerodynamic, propulsive and gravitational). The overall simulation has been completed by means of the PIPER PA–30 propulsion system description as well as the models of atmosphere, servo–actuators and input–output sensors. The description of the Navigation, Guidance and Control (NGC) system has been also included.

The PM residual generators have been designed on the basis of the linearised aircraft simulation model in different flight condition. Since the aircraft simulation model does not match the hypothesis to apply the NLGA methodology, a simplified nonlinear model has been developed for the purpose of the NLGA–based filters design.

The final performances have been evaluated by adopting a typical aircraft reference trajectory embedding several steady–state flight conditions, such as straight flight phases and coordinated turns. Comparisons with different disturbance decoupling methods for FDI based on Neural Networks (NN) and Unknown Input Kalman Filter (UIKF) have been also provided. Finally, extensive experiments exploiting Monte–Carlo analysis are used for assessing the overall capabilities of the developed FDI methods, in the presence of uncertainty, measurement and modelling errors.

The thesis is organised as follows. A description of the aircraft simulation model is provided in Chapter 2. Chapter 3 presents the design details for the proposed PM scheme, whilst in Chapter 4 a description of the NLGA–based design schemes is reported. In Chapter 5 the effectiveness of the proposed residual generators applied to the aircraft simulator and extensive simulation results are reported. Finally, Chapter 6 summarises the contributions and the achievements of the thesis (Benini *et al.* 2008*a*, Castaldi *et al.* 2009, Beghelli *et al.* 2007*a*, Beghelli *et al.* 2007*b*, Simani and Benini 2007, Benini *et al.* 2008*b*, Bonfè *et al.* 2008, Benini *et al.* 2009, Bonfè *et al.* 2006, Bonfè *et al.* 2007*b*, Bonfè *et al.* 2007*a*).

For the readers not familiar with the basic principles of fault diagnosis, a review of model–based FDI is reported in Appendix A (Simani *et al.* 2002, Chen and Patton 1999).

Chapter 2

Aircraft Simulation Model

This chapter provides a description of the PIPER PA–30 aircraft simulation model. The 6 DoF aircraft model is derived in Section 2.1. The description of the overall simulation model is completed in Section 2.2. Finally, the mathematical models used for the FDI purpose are developed in Section 2.3.

2.1 6 DoF Aircraft Model

The aircraft can be considered as a rigid body with a given mass and moments of inertia. For a rigid body, the system undergoes no deformation and should possess only 6 degrees of freedom, namely 3 translations and 3 rotations.

The following axes systems are considered:

- An Earth–fixed axes system $OXYZ$, such that the plane (X, Y) coincides with the Earth surface at the sea level and the axis Z represents the aircraft altitude H changed of sign, *i.e.* $H = -Z$. This axes system is assumed to be an inertial frame of reference.

- A body–fixed axes system $O'xyz$ (the so–called body axes), whose origin O' is located identically at the aircraft center gravity C. For such a system, the axis x points forward out of the nose of the aircraft; the axis y points out through the right wing; and the axis z points down.

The motion of the aircraft can be described by:

1. Translation of the origin O' of the body axes.

2. Rotation of the axes with respect to the inertial space.

2.1.1 Force Equations

Let us consider Newton laws applying to the linear momentum

$$\left(\frac{dp}{dt}\right)_{OXYZ} = \left(\frac{dp}{dt}\right)_{O'xyz} + \omega \times p = F \tag{2.1}$$

where $F = [F_x \; F_y \; F_z]^{\mathrm{T}}$ represents the external forces applied to the body and the linear momentum is defined as

$$p = m\, V_C \tag{2.2}$$

where m is the total body mass and V_C is the velocity of the center of mass. Hence (2.1) becomes

$$m\,(\dot{V}_C + \omega \times V_C) = F \tag{2.3}$$

Let us point out the components along the body axes of the linear velocity V_C and angular velocity ω

$$V_C = \begin{bmatrix} u \\ v \\ w \end{bmatrix} \qquad \omega = \begin{bmatrix} p_\omega \\ q_\omega \\ r_\omega \end{bmatrix} \tag{2.4}$$

where p_ω, q_ω and r_ω are the roll, pitch and yaw rate, respectively. Then the force equations of motion along the body axes are given by

$$\begin{aligned} m\,(\dot{u} - r_\omega\, v + q_\omega\, w) &= F_x \\ m\,(\dot{v} - p_\omega\, w + r_\omega\, u) &= F_y \\ m\,(\dot{w} - q_\omega\, u + p_\omega\, v) &= F_z \end{aligned} \tag{2.5}$$

where the force components F_x, F_y and F_z on the right–hand side of the above equations are due to gravitational, aerodynamic and thrust forces.

2.1.2 Moment Equations

Let us consider Newton laws applying to the angular momentum

$$\left(\frac{dH_C}{dt}\right)_{OXYZ} = \left(\frac{dH_C}{dt}\right)_{O'xyz} + \omega \times H_C = M \tag{2.6}$$

where $M = [M_x \; M_y \; M_z]^{\mathrm{T}}$ represents the external moments applied to the body and the angular momentum is defined as

$$H_C = I\, \omega \tag{2.7}$$

where

$$I = \begin{bmatrix} I_x & 0 & -I_{xz} \\ 0 & I_y & 0 \\ -I_{xz} & 0 & I_z \end{bmatrix} \tag{2.8}$$

is the inertia moment matrix of the body. Note that the form of I is due to the symmetry properties of the considered aircraft. Hence (2.6) becomes

$$I\,\dot{\omega} + \omega \times I\,\omega = M \tag{2.9}$$

Using the above definitions for ω and I, the moment equations of motion can be written about the body axes in the following way

$$\begin{aligned} I_x\,\dot{p}_\omega - (I_y - I_z)\,q_\omega\,r_\omega - I_{xz}\,(\dot{r}_\omega + p_\omega\,q_\omega) &= M_x \\ I_y\,\dot{q}_\omega - (I_z - I_x)\,r_\omega\,p_\omega - I_{xz}\,(r_\omega^2 - p_\omega^2) &= M_y \\ I_z\,\dot{r}_\omega - (I_x - I_y)\,p_\omega\,q_\omega - I_{xz}\,(\dot{p}_\omega - q_\omega\,r_\omega) &= M_z \end{aligned} \tag{2.10}$$

where the moments components M_x, M_y and M_z on the right side of the above equations are due to aerodynamic and propulsion forces. Note that there is no contribution from the gravitational force since these moments are taken about the center of gravity.

2.1.3 Euler Angles

The angular velocity components p_ω, q_ω and r_ω about the body axes x, y and z cannot be integrated to obtain the corresponding angular displacements about these axes. In other words, the orientation of the aircraft in space is not known until we describe the three rotational degrees of freedom in terms of a set of independent coordinates. Of course, such a set is not necessarily unique. One useful set of angular displacements, the so–called Euler angles, is obtained through successive rotations about three (not necessarily orthogonal) axes as follows.

Let us start with a set of inertial axes $OXYZ$ and perform the following rotations in a particular order:

1. Rotation about the Z–axis (*i.e.* yaw) through an angle ψ. This rotation leads to the new coordinates system (x_1, y_1, z_1).

2. Rotation about the y_1–axis (*i.e.* pitch) through an angle θ. This rotation leads to the new coordinates system (x_2, y_2, z_2).

3. Rotation about the x_2-axis (*i.e.* roll) through an angle ϕ. This rotation leads to the new coordinates system (x_3, y_3, z_3).

The Euler angles for an aircraft are defined as above in terms of ψ, θ and ϕ. Those angles are also known as the heading, elevation and bank angle, respectively. At each rotation, components of a vector expressed in the coordinate frame before and after the rotation are related through a rotation matrix. Namely:

1. ψ rotation

$$\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = R_z(\psi) \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{2.11}$$

2. θ rotation

$$\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = R_y(\theta) \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} \tag{2.12}$$

3. ϕ rotation

$$\begin{bmatrix} x_3 \\ y_3 \\ z_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & \sin\phi \\ 0 & -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = R_x(\phi) \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} \tag{2.13}$$

It is worth observing that the rotation matrices defined above are orthogonal, hence nonsingular and invertible.

The angular velocity ω can be expressed as a function of the Euler angles in the following way

$$\begin{bmatrix} p_\omega \\ q_\omega \\ r_\omega \end{bmatrix} = R_x(\phi)\, R_y(\theta) \begin{bmatrix} 0 \\ 0 \\ \dot{\psi} \end{bmatrix} + R_x(\phi) \begin{bmatrix} 0 \\ \dot{\theta} \\ 0 \end{bmatrix} + \begin{bmatrix} \dot{\phi} \\ 0 \\ 0 \end{bmatrix} \tag{2.14}$$

that is

$$\begin{aligned} p_\omega &= \dot{\phi} - \dot{\psi}\, \sin\theta \\ q_\omega &= \dot{\theta}\, \cos\phi + \dot{\psi} \cos\theta\, \sin\phi \\ r_\omega &= \dot{\psi}\, \cos\theta\, \cos\phi - \dot{\theta}\, \sin\phi \end{aligned} \tag{2.15}$$

Since a flat–Earth model is considered, the gravitational force is always pointed along the Z–axis of the inertial frame of reference. Hence the components of the gravitational forces along the body axes are obtained as follows

$$\begin{bmatrix} F_{GRAx} \\ F_{GRAy} \\ F_{GRAz} \end{bmatrix} = R_x(\phi)\, R_y(\theta) \begin{bmatrix} 0 \\ 0 \\ m\, g(H) \end{bmatrix} \tag{2.16}$$

that is

$$\begin{aligned} F_{GRAx} &= -m\, g(H)\, \sin\theta \\ F_{GRAy} &= m\, g(H)\, \cos\theta\, \sin\phi \\ F_{GRAz} &= m\, g(H)\, \cos\theta\, \cos\phi \end{aligned} \tag{2.17}$$

where $g(H)$ represents the gravity acceleration at current altitude.

The rotational matrices defined above can be also exploited to determine the aircraft position (in the inertial space) in terms of its linear velocity components u, v and w in the body-fixed axes

$$\begin{bmatrix} \dot{X} \\ \dot{Y} \\ \dot{Z} \end{bmatrix} = R_z(\psi)^{\mathrm{T}}\, R_y(\theta)^{\mathrm{T}} R_x(\phi)^{\mathrm{T}} \begin{bmatrix} u \\ v \\ w \end{bmatrix} \tag{2.18}$$

that is

$$\begin{aligned}
\dot{X} &= u\,\cos\psi\,\cos\theta + v\,(-\sin\psi\,\cos\theta + \cos\psi\,\sin\theta\,\sin\phi) \\
&\quad + w\,(\sin\psi\,\sin\phi + \cos\psi\,\sin\theta\,\cos\phi) \\
\dot{Y} &= u\,\sin\psi\,\cos\theta + v\,(\cos\psi\,\cos\phi + \sin\psi\,\sin\theta\,\sin\phi) \\
&\quad + w\,(-\cos\psi\,\sin\phi + \sin\psi\,\sin\theta\,\cos\phi) \\
\dot{Z} &= -u\,\sin\theta + v\,\cos\theta\,\sin\phi + w\,\cos\theta\,\cos\phi
\end{aligned} \tag{2.19}$$

2.1.4 True Air Speed and Aerodynamic Angles

Major contributions to the forces and moments in a flight vehicle are coming from the aerodynamics of wings, body and tail surfaces. It would be difficult to express these in terms of the aircraft motion variables u, v and w. However it is much easier to express them in terms of the true air speed V, angle of attack α and angle of sideslip β.

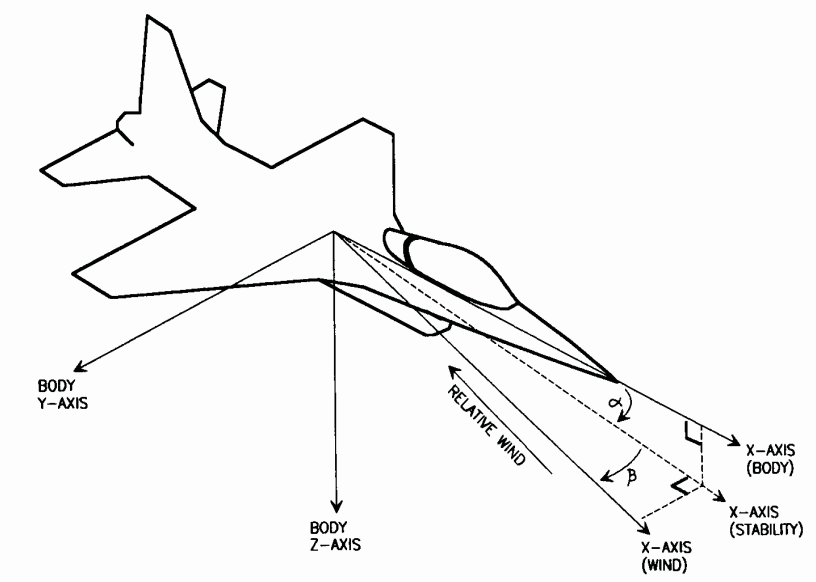


Figure 2.1: Aircraft axes and angles.

The true air speed is the speed of an aircraft relative to the airmass in which it flies, *i.e.* the magnitude of the vector difference of the velocity of the aircraft and the velocity of the air. The angles of attack and sideslip (said aerodynamic angles) are defined by performing a plane rotation about the body y–axis, followed by a plane rotation about the new z–axis, such that the final x–axis is aligned directly into the relative wind (*i.e.* the direction of the air over the aircraft wings and control surfaces). The first rotation defines the stability axes, and the angle of attack is the angle between the body–fixed x–axis and the stability x–axis. The second rotation leads to a set of wind axes, and the sideslip angle is the angle between the stability x–axis and the wind x–axis (see Figure 2.1).

The linear velocity components u, v and w can be expressed in terms of V, α and β as follows

$$\begin{aligned} u &= V\ \cos\beta\ \cos\alpha \\ v &= V\ \sin\beta \\ w &= V\ \cos\beta\ \sin\alpha \end{aligned} \tag{2.20}$$

Note that by substituting (2.20) into (2.5) the following equations are obtained

$$\begin{aligned} \dot{u} &= \frac{F_x}{m} + r_\omega\,(V\ \sin\beta) - q_\omega\,(V\ \cos\beta\ \sin\alpha) \\ \dot{v} &= \frac{F_y}{m} + p_\omega\,(V\ \cos\beta\ \sin\alpha) - r_\omega\,(V\ \cos\beta\ \cos\alpha) \\ \dot{w} &= \frac{F_z}{m} + q_\omega\,(V\ \cos\beta\ \cos\alpha) - p_\omega\,(V\ \sin\beta) \end{aligned} \tag{2.21}$$

By differentiating the equations (2.20) with respect to time also the linear acceleration components $\dot{u}$, $\dot{v}$ and $\dot{w}$ can be derived in terms of V, α and β

$$\begin{bmatrix} \dot{u} \\ \dot{v} \\ \dot{w} \end{bmatrix} = \begin{bmatrix} \cos\alpha & -V\ \sin\alpha\ \cos\beta & -V\ \cos\alpha\ \sin\beta \\ \sin\beta & 0 & V\ \cos\beta \\ \sin\alpha\ \cos\beta & V\ \cos\alpha\ \cos\beta & -V\ \sin\alpha\ \sin\beta \end{bmatrix} \begin{bmatrix} \dot{V} \\ \dot{\alpha} \\ \dot{\beta} \end{bmatrix} \tag{2.22}$$

Solving for $\dot{V}$, $\dot{\alpha}$ and $\dot{\beta}$, the following linear system is obtained

$$\begin{bmatrix} \dot{V} \\ \dot{\alpha} \\ \dot{\beta} \end{bmatrix} = \begin{bmatrix} \cos\alpha\ \cos\beta & \sin\beta & \sin\alpha\ \cos\beta \\ -\sin\alpha/(V\ \cos\beta) & 0 & \cos\alpha/(V\ \cos\beta) \\ -\cos\alpha\ \cos\beta/V & \cos\beta/V & -\sin\alpha\ \sin\beta/V \end{bmatrix} \begin{bmatrix} \dot{u} \\ \dot{v} \\ \dot{w} \end{bmatrix} \tag{2.23}$$

2.1.5 Overall Model

The equations governing the motion of a rigid body aircraft are summarised in the following.

- Equations representing the time derivative of the linear momentum related to total forces applied to the aircraft (obtained substituting (2.21) into (2.23))

$$\begin{aligned} \dot{V} &= F_x\,\frac{\cos\alpha\ \cos\beta}{m} + F_y\,\frac{\sin\beta}{m} + F_z\,\frac{\sin\alpha\ \cos\beta}{m} \\ \dot{\alpha} &= \frac{-F_x\ \sin\alpha + F_z\ \cos\alpha}{m\,V\ \cos\beta} + q_\omega - \left(p_\omega\ \cos\alpha + r_\omega\ \sin\alpha\right)\ \tan\beta \\ \dot{\beta} &= \frac{-F_x\ \cos\alpha\ \sin\beta + F_y\ \cos\beta - F_z\ \sin\alpha\ \sin\beta}{m\,V} + p_\omega\ \sin\alpha - r_\omega\ \cos\alpha \end{aligned} \tag{2.24}$$

- Equations representing the time derivative of the the angular momentum related to total moments applied to the aircraft (obtained from (2.10))

$$\begin{aligned}
\dot{p}_\omega &= \frac{M_x\, I_z + M_z\, I_{xz} + p_\omega\, q_\omega\, I_{xz}\left(I_x - I_y + I_z\right) + q_\omega\, r_\omega\left(I_y\, I_z - {I_{xz}}^2 - {I_z}^2\right)}{I_x\, I_z - {I_{xz}}^2} \\
\dot{q}_\omega &= \frac{M_y + p_\omega\, r_\omega\left(I_z - I_x\right) - p_\omega^2\, I_{xz} + r_\omega^2\, I_{xz}}{I_y} \\
\dot{r}_\omega &= \frac{M_x\, I_{xz} + M_z\, I_x + p_\omega\, q_\omega\left({I_x}^2 - I_x\, I_y + {I_{xz}}^2\right) + q_\omega\, r_\omega\, I_{xz}\left(-I_x + I_y - I_z\right)}{I_x\, I_z - {I_{xz}}^2}
\end{aligned} \tag{2.25}$$

- Equations representing the cinematic equations for the Euler angles propagation (obtained from (2.15))

$$\begin{aligned}
\dot{\phi} &= p_\omega + q_\omega\, \sin\phi\, \tan\theta + r_\omega\, \cos\phi\, \tan\theta \\
\dot{\theta} &= q_\omega\, \cos\phi - r_\omega\, \sin\phi \\
\dot{\psi} &= \frac{q_\omega\, \sin\phi + r_\omega\, \cos\phi}{\cos\theta}
\end{aligned} \tag{2.26}$$

- Equations relating the true air speed to the position coordinates respect to an inertial reference frame with the origin at the sea level (obtained substituting (2.20) into (2.19))

$$\begin{aligned}
\dot{X} &= V\, \cos\psi\left[\cos\alpha\, \cos\beta\, \cos\theta + \sin\theta\left(\sin\beta\, \sin\phi + \sin\alpha\, \cos\beta\, \cos\phi\right)\right] \\
&\quad - V\, \sin\psi\left(\sin\beta\, \cos\phi - \sin\alpha\, \cos\beta\, \sin\phi\right) + V_{Ax} \\
\dot{Y} &= V\, \sin\psi\left[\cos\alpha\, \cos\beta\, \cos\theta + \sin\theta\left(\sin\beta\, \sin\phi + \sin\alpha\, \cos\beta\, \cos\phi\right)\right] \\
&\quad + V\, \cos\psi\left(\sin\beta\, \cos\phi - \sin\alpha\, \cos\beta\, \sin\phi\right) + V_{Ay} \\
\dot{H} &= V\, \cos\alpha\, \cos\beta\, \sin\theta - V\, \cos\theta\left(\sin\beta\, \sin\phi + \sin\alpha\, \cos\beta\, \cos\phi\right) - V_{Az}
\end{aligned} \tag{2.27}$$

Total force and moment components can be expressed by the combinations of aerodynamic, thrust and gravitational contribution as follows

$$\begin{aligned}
F_x &= F_{AERx} + T_h - m\, g(H)\, \sin\theta \\
F_y &= F_{AERy} + m\, g(H)\, \cos\theta\, \sin\phi \\
F_z &= F_{AERz} + m\, g(H)\, \cos\theta\, \cos\phi
\end{aligned} \tag{2.28}$$

$$\begin{aligned}
M_x &= M_{AERx} \\
M_y &= M_{AERy} + d_T\, T_h \\
M_z &= M_{AERz}
\end{aligned} \tag{2.29}$$

where T_h is the thrust and d_T is the distance between the aircraft center of gravity and the thrust axis. Note that the gravitational contribution to total forces is obtained from (2.17). Note also that there is not gravitational contribution to the total moments.

As to the aerodynamic forces $F_{AER(.)}$ and moments $M_{AER(.)}$, a set of local approximations has been computed and scheduled depending on the values assumed by true air speed, flap, altitude, curve radius, and flight path angle (*i.e.* the angle between velocity vector respect to air and its projection over the horizontal plane). In this way, it is possible to obtain a mathematical model for each flight condition. This model is suitable suitable for a state–space representation, as it can be made explicit.

The parameters in the analytic representation of the aerodynamic actions have been obtained from wind tunnel experimental data, as reported in (Fink and Freeman 1969, Koziol 1971), and the aerodynamic actions are expressed along the axes of the wind reference system. It should be observed that aerodynamic forces and moments are not implemented by the classical linearised expressions (stability derivatives) as reported in Flight Dynamic textbook, (Etkin and Reid 1996). Aerodynamic actions, in fact, are implemented by means of cubic splines approximating the non–linear experimental curves given in (Fink and Freeman 1969).

Remark 1. *The thrust term T_h depends on the throttle aperture percentage δ_{th} (see Section 2.2.1), whilst the aerodynamic action terms $F_{AER(.)}$ and $M_{AER(.)}$ depends on the control surfaces deflection angles,* i.e. *δ_a, δ_e and δ_r that are the aileron, elevator and rudder deflection angles, respectively. δ_{th}, δ_a, δ_e and δ_r represent the control inputs of the monitored system for FDI purpose.*

2.2 Simulation Model Subsystems

2.2.1 Engine Model

A first order dynamic model of a 4-pistons aspirated engine with the throttle aperture as input and the thrust intensity as output has been considered. The propulsion system of the PIPER PA–30 aircraft consists of two engines of this type.

The main advantage of this model consists in the fact that it is both simple and based on the dynamic balance of the torques insisting on the propeller. The engine model can be written as follows

$$\dot{n_e} = \frac{(1 - \eta_{pr} - \eta_{air})}{I_{pr}} \left(\frac{60}{2\,\pi}\right)^2 \frac{\rho(H)}{\rho(0)} \sqrt{\frac{T(H)}{T(0)}}\, \frac{\delta_{th}}{n_e}\, P_c(n_e) - \frac{J_v}{I_{pr}} \left(\frac{2\,\pi}{60}\right)^2 n_e^3 \tag{2.30}$$

with

$$T_h = \frac{2\,\eta_{pr}}{V\,\cos\alpha\,\cos\beta}\, \frac{\rho(H)}{\rho(0)} \sqrt{\frac{T(H)}{T(0)}}\, \delta_{th}\, P_c(n_e) \tag{2.31}$$

where n_e is the engine shaft angular rate, J_v is the viscous friction coefficient of the transmission shaft, I_{pr} is the propeller moment of inertia, η_{pr} is the propeller efficiency, η_{air} is the percentage loss of available power due to air, $\rho(H)$ is the air density at current altitude, $T(H)$ is the air temperature at current altitude and $P_c(n_e)$ is the engine power behaviour with respect to n_e at full throttle.

The model (2.30) is obtained by the equilibrium of torques (inertial, viscous friction, load and driving torque T_d) applied to the engine shaft with the assumption of a propeller with constant efficiency

$$I_{pr}\,\frac{2\,\pi}{60}\,\dot{n_e} + J_v\,\left(\frac{2\,\pi}{60}\,n_e\right)^3 + (\eta_{pr} + \eta_{air})\,T_d = T_d \tag{2.32}$$

with

$$T_d = \frac{60}{2\,\pi}\,\frac{\mathrm{BP}(H)}{n_e} \tag{2.33}$$

where $\mathrm{BP}(H)$ is the brake power at current altitude (Ojha 1995)

$$\mathrm{BP}(H) = \frac{\rho(H)}{\rho(0)}\,\sqrt{\frac{T(H)}{T(0)}}\,\mathrm{BP}(0) \qquad \mathrm{BP}(0) = \delta_{th}\,P_c(n_e) \tag{2.34}$$

The nonlinear curve $P_c(n_e)$ has been approximated by means of a cubic spline derived from (Koziol 1971).

2.2.2 Atmosphere Model

Air Temperature and Density

The atmoshpere model describes the behaviour of temperature and air density as a function of altitude above the mean sea level.

The temperature is considered a linear decreasing function of altitude with a constant slope $G_T = 6.5\,^{\circ}\mathrm{K/Km}$ up to a maximum altitude of 11 Km, starting with a value of $T(0)$ at the sea level.

The air is assumed to be a perfect gas, therefore the air density is related to the altitude by the following differential equation

$$\frac{d\rho}{dH} = -\rho\,\frac{g(H)\,M}{R\,T(H)} \tag{2.35}$$

with

$$g(H) = g(0)\,\left(\frac{r}{r+H}\right)^2 \tag{2.36}$$

where M is the molar mass of the air mixture, R is the universal constant of perfect gases and r is the mean earth radius. Solving the differential equation the following air density model is obtained

$$\rho(H) = \rho(0)\,\left[\frac{T(0)}{r}\,\left(\frac{r+H}{T(0) - G_T\,H}\right)\right]^{\frac{K_\rho\,G_T}{(r\,G_T + T(0))^2}}\,e^{\frac{K_\rho\,H}{r\,(r\,G_T + T(0))\,(r+H)}} \tag{2.37}$$

where

$$K_\rho = -\frac{M\,g(0)\,r^2}{R} \tag{2.38}$$

Wind, Wind Shear and Wind Gusts

The atmosphere model embeds also a mathematical model description of the wind, wind shear and wind gusts.

The wind is modeled as a constant velocity bias vector (whose components are V_{Ax}, V_{Ay} and V_{Az}) of the atmosphere respect to the ground.

The wind shear is a vertical gradient of the wind velocity. Its effects are relevant for low altitude and it can be described by equations that represent a good approximation of the wind shear model published in (Moorhouse and Woodcock 1980) by means of the following smooth functions

$$\begin{aligned} V_{Ax} &= \cos(\psi_{wind})\, O_{sat} \left(1 - e^{-\frac{5\,H}{H_{lim}}}\right) \\ V_{Ay} &= \sin(\psi_{wind})\, O_{sat} \left(1 - e^{-\frac{5\,H}{H_{lim}}}\right) \\ V_{Az} &= V_{sat} \left(1 - e^{-\frac{5\,H}{H_{lim}}}\right) \end{aligned} \tag{2.39}$$

where ψ_{wind} is the direction of the arrival of the wind, O_{sat} is the wind maximum horizontal ground speed, V_{sat} is the wind maximum vertical ground speed and H_{lim} is the reference maximum altitude for wind shear. A suitable value for the reference maximum altitude is $H_{lim} = 60\,\mathrm{m}$. The wind shear velocity gradient effect can be assimilated to a motion in a non inertial reference frame and therefore causes the so–called apparent forces, extremely dangerous during the approach phase.

While the wind consist in the atmosphere steady motion, the wind gusts represent an air motion with zero mean velocity. Wind gusts are modeled as body axes air velocity (w_u, w_v and w_w) described by means of colored stochastic processes generated by first order shaping filters with the correlation times and wind covariance (Moorhouse and Woodcock 1980) specified in Table 2.1

Table 2.1: Wind gusts model parameters.

Correlation time	**Wind covariance**
$\tau_u = 2.326\,\mathrm{s}$	$E[w_u^2] = 0.7\,(\mathrm{m/s})^2$
$\tau_v = 7.143\,\mathrm{s}$	$E[w_v^2] = 0.7\,(\mathrm{m/s})^2$
$\tau_w = 0.943\,\mathrm{s}$	$E[w_w^2] = 0.7\,(\mathrm{m/s})^2$

Remark 2. *The wind gusts represent the disturbances acting on the system. In the residual generators design those disturbances must be decoupled in order to assure the robustness of the proposed FDI techniques.*

2.2.3 Servo–Actuators Model

The main task of the servo–actuators is to move the control surfaces: elevator, aileron and rudder. Moreover, there is a forth servo–actuator that steers the throttle positioning.

In the considered aircraft the servo–actuators are DC–motors. Therefore they are modeled as second order dynamic systems without zeros. In order to avoid out of range of the deflection surfaces, overshoots during transient responses are unwished. Consequently, the loop controls of the actuators have been designed with gain constants assuring real and coincident poles to the servos. The values of the poles of the transfer functions used for each servo are shown in Table 2.2.

Table 2.2: Transfer function poles of the servos.

Elevator servo	**Aileron servo**	**Rudder servo**	**Throttle servo**
$-3.45\,\mathrm{s}^{-1}$	$-3.45\,\mathrm{s}^{-1}$	$-3.45\,\mathrm{s}^{-1}$	$-8.26\,\mathrm{s}^{-1}$

2.2.4 Measurement Errors Description

In the following, a brief description of the measurement subsystems used by the simulation model is provided. It is worth noting that the sensor models embed all the possible sources of disturbance (calibration and alignment errors, scale factor, white and coloured noises, limited bandwidth, g–sensitivity, gyro drift, etc.).

Command Surfaces Deflection Measurements

It is assumed that the deflection angles δ_e, δ_a, δ_r and δ_{th} are acquired with a sample rate of 100 Hz by means of potentiometers. These sensors are affected by errors modelled by two additive components: bias and white noise. The bias values and the standard deviation (std) of the noises are given in Table 2.3. The reported parameters have been obtained by means of experimental tests performed at the aerospace engineering laboratory of the University of Bologna.

Table 2.3: Input sensor errors parameters.

Input sensor	**Bias**	**White Noise Std**
Elevator deflection angle	0.0052 rad	0.0053 rad
Aileron deflection angle	0.0052 rad	0.0053 rad
Rudder deflection angle	0.0052 rad	0.0053 rad
Throttle aperture	1%	1%

Angular Rate Measurement

It is assumed that the angular rate measurements are given by a set of three gyroscopes of an Inertial Measurement Unit (IMU) with a sample rate of 100 Hz. The errors affecting this measurement unit can be classified as follows (Randle and Horton 1997):

- Errors due to non–unitary scale factor, modelled by a multiplicative factor belonging to the range $[0.99, 1.01]$.
- Alignment error of spin axes with respect to body (reference) axes. These errors can be modelled by considering each spin axis oriented in a 3D space by means of an azimuth and elevation angle with respect to its reference axis. It this way, the alignment errors can be described by six error angles up to 1^o. It is worth observing that the errors previously considered are generated by means of uniform random variables updated every simulation.
- Limited bandwidth of the considered gyro (10 Hz).
- g–sensitivity ($72^o/(\mathrm{h\,g})$).
- Additive white noise ($216^o/\mathrm{h}$).
- Gyro drift, described by a coloured stochastic process characterised by a standard deviation of $1080^o/\mathrm{h}$ and a decay time of 20 min.

Attitude Angle Measurement

The angles are actually generated by a digital filtering system based on a DSP that processes both the angular rate and the accelerations provided by the IMU with a sample rate of 100 Hz.

The angle generation system has been considered equivalent to a mechanical vertical gyro for aeronautical purposes (artificial horizon). As reported in ((Bryson 1994), Chapter 11), the measurement errors are due to the sum of two causes:

- A systematic error generated by the apparent vertical. This effect cannot be neglected because the fault diagnosis, as it will be shown in the following, has to be performed in coordinated turn flight condition.
- A white noise modelling the imperfection of both the system and the environment influences.

The behaviour of this angle measurement system is such that the previous two effects are correlated by a first order filter system with time constant equal to 60 s (Bryson 1994). Therefore, the resulting attitude angle measurements are affected by an additive coloured noise characterised by a standard deviation of 1^o.

The angular rate measurements exploited by the attitude angle estimation system are provided by a gyroscope unit that is different from the gyroscope device estimating directly the angular rates. In fact, the gyroscope unit adopted for attitude angle estimation must guarantee a low drift, since the angular rate signals measured on this unit are integrated by the system to obtain angles. On the other hand, the gyroscope device directly providing the angular rate measurements requires larger bandwidths (Titterton and Weston 2005).

Air Data System (ADS)

It is assumed that the ADS unit consists of an Air Data Computer (ADC) providing measurements with a sample rate of 1 Hz. The errors affecting the true air speed can be classified as follows:

- Calibration error affecting the differential pressure sensor. This error leads to a true air speed computation systematic error, performed the ADC, fulfilling the ARINC (Aeronautical Radio Inc.) (ARINC 1998) accuracy requirements (2 m/s) (Bryson 1994).
- Additive coloured noise induced by wind gusts and atmospheric turbulence (std 1 m/s and correlation time 2.3 s).
- Additive white noise (std 0.5 m/s) modelling the imperfection of the system and the environment influences.

With regards to the altitude, errors can be classified as:

- Calibration error affecting the static pressure sensor. This error leads to an altitude computation systematic error, performed the ADC, fulfilling the ARINC accuracy requirements (5 m) (ARINC 1998).
- Additive white noise (std 1 m) modelling the imperfection of the system and the environment influences.

With regards to the attack and sideslip angle, errors can be classified as:

- Calibration error affecting the wing boom sensors. This systematic error is 1^o for both angles.
- Additive white noise (std 2^o) modelling the imperfection of the sensor and the wind turbulence effects.

Heading Reference System (HRS)

This unit is assumed to consist of a magnetic compass coupled to a directional gyro. As reported in (Bryson 1994) the measurement errors are due to the sum of two causes:

- A systematic error generated by a bias of the magnetic compass (1^o).
- A white noise modelling the imperfection of the system and the environment influences.

The behaviour of the HRS system is such that the two previous effects are correlated by a first order filter with time constant equal to 60 s (Bryson 1994). Hence, the resulting heading measurement is affected by an additive coloured noise characterised by a std 1^o.

Similar assumptions regarding the attitude angle and angular rate estimation hold for the HRS system, where the directional gyro unit is different from the other measurement subsystem components.

Engine Shaft Rate Measurement

The engine shaft rate is measured by means of an incremental encoder whose errors are modelled as a white noise. The quantisation error of the encoder is determined by a resolution of 10000 pulse/rev.

2.2.5 NGC System

In Figure 2.2 the overall architecture is shown. The blocks corresponding to the navigation, guidance and control functions are highlighted with the processed information.

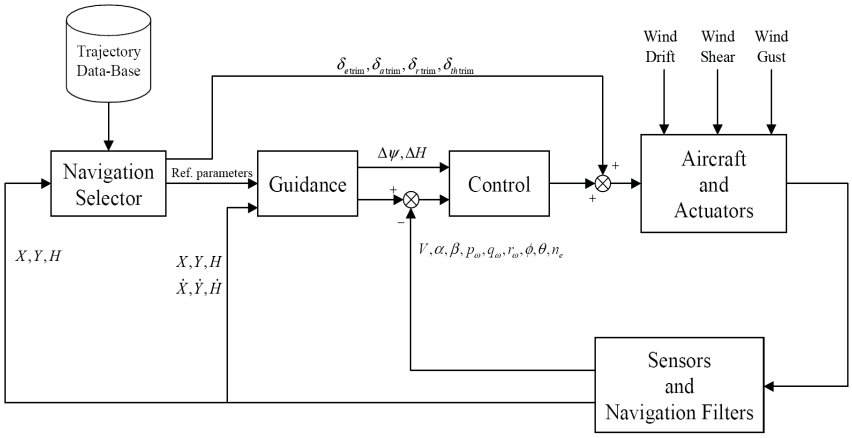


Figure 2.2: Overall architecture of the NGC system.

Navigation System

The aim of the navigation system is twofold:

1. To estimate the aircraft state, that is position, velocity and attitude.
2. To select the trajectory branch to be followed and to provide its parameters to downstream blocks.

It is composed by three subsystems: the sensors and navigation filters, the navigation selector and the trajectory data–base.

As to the first task, usually the estimate of the aircraft state is accomplished by means of a data fusion, performed inside the sensors and navigation filters subsystem, that processes the signals provided by the aircraft sensors: GPS, barometric altimeter, Pitot tube, attitude and heading reference system, rate gyros.

The second task is performed by the navigation selector subsystem that interacts with the trajectory data–base subsystem. Therefore the data-base has to contain the characteristic parameters which describe the following trajectory branches:

- The class of branches corresponding to leveled wing, straight and symmetric flight conditions.
- The class of branches corresponding to horizontal coordinated turns.

These classes of trajectory branches correspond to standard steady flight conditions, so that it is straightforward to determine the trim values for the control surfaces deflection, throttle aperture, attitude angles, aerodynamic angles, angular rates and engine rpm.

Guidance System

The main task of the guidance system is to provide to the control block:

- The error on the velocity vector direction ($\Delta\psi$, ΔH) on the basis of the actual values of inertial position and velocity.
- The reference values of true air speed (V), aerodynamic angles (α, β), inertial angular rates (p_ω, q_ω, r_ω), attitude angles (ϕ, θ) and engine angular rate (n_e) directly from the navigation selector.

Control System

The control system stabilises the aircraft around the selected stationary flight condition. It is projected by means of classical LQ optimal law applied to attitude linearised models.

2.3 Aircraft FDI Model

This section describes the so–called aircraft FDI model, *i.e.* the model used to design the residual generators, for both the PM and the NLGA–based techniques.

2.3.1 PM FDI Model

The proposed PM FDI scheme can be properly applied to a linear system. Hence the aircraft simulation model presented in the previous section has to be linearised for different flight condition. The linear model embeds the linearisation of both the 6–DoF model and the propulsion system as follows

$$\dot{x}(t) = A\,x(t) + B\,c(t) + E\,d(t) \tag{2.40}$$

with

$$\begin{aligned} x(t) &= \left[\ \Delta V(t) \quad \Delta\alpha(t) \quad \Delta\beta(t) \quad \Delta p_\omega(t) \quad \Delta q_\omega(t) \quad \Delta r_\omega(t) \right. \\ &\qquad \left. \ldots \quad \Delta\phi(t) \quad \Delta\theta(t) \quad \Delta\psi(t) \quad \Delta H(t) \quad \Delta n_e(t)\ \right]^{\mathrm{T}} \\ c(t) &= \left[\ \Delta\delta_e(t) \quad \Delta\delta_a(t) \quad \Delta\delta_r(t) \quad \Delta\delta_{th}(t)\ \right]^{\mathrm{T}} \\ d(t) &= \left[\ w_u(t) \quad w_v(t) \quad w_w(t)\ \right]^{\mathrm{T}} \end{aligned} \tag{2.41}$$

where Δ denotes the variations of the considered variables, while $c(t)$ and $d(t)$ are the control inputs and the disturbances respectively. The disturbance contribution of the wind gusts as air velocity components, w_u, w_v and w_w, along body axes was also considered. The output equation associated with the model (2.40) is of the type $y(t) = C\,x(t)$, where the rows of C correspond to rows of the identity matrix, depending on the measured variables.

2.3.2 NLGA FDI Model

The NLGA FDI scheme requires a nonlinear input affine system (De Persis and Isidori 2001), but the adopted simulation model of the aircraft does not fulfil this requirement. For this reason, the following simplified aircraft model is used

$$
\begin{aligned}
\dot{V} &= -\frac{(C_{D0} + C_{D\alpha}\alpha + C_{D\alpha_2}\alpha^2)}{m} V^2 + g\,(\sin\alpha\cos\theta\cos\phi - \cos\alpha\sin\theta) \\
&\quad + \frac{\cos\alpha}{m}\frac{t_p}{V}\,(t_0 + t_1 n_e)\,\delta_{th} + w_v \sin\alpha \\
\dot{\alpha} &= -\frac{(C_{L0} + C_{L\alpha}\alpha)}{m} V + \frac{g}{V}\,(\cos\alpha\cos\theta\cos\phi + \sin\alpha\sin\theta) + q_\omega \\
&\quad - \frac{\sin\alpha}{m}\frac{t_p}{V^2}\,(t_0 + t_1 n_e)\,\delta_{th} + \frac{\cos\alpha}{V} w_v \\
\dot{\beta} &= \frac{(C_{D0} + C_{D\alpha}\alpha + C_{D\alpha_2}\alpha^2)\sin\beta + C_{Y\,\beta}\beta\cos\beta}{m} V + g\frac{\cos\theta\sin\phi}{V} \\
&\quad + p_\omega \sin\alpha - r_\omega\cos\alpha - \frac{\cos\alpha\sin\beta}{m}\frac{t_p}{V^2}\,(t_0 + t_1 n_e)\,\delta_{th} + \frac{1}{V} w_\ell \\
\dot{p}_\omega &= \frac{(C_{l\beta}\beta + C_{lp}\,p_\omega)}{I_x} V^2 \;+\; \frac{(I_y - I_z)}{I_x}\,q_\omega r_\omega + \frac{C_{\delta_a}}{I_x} V^2 \delta_a \\
\dot{q}_\omega &= \frac{(C_{m0} + C_{m\alpha}\alpha + C_{mq}\,q_\omega)}{I_y} V^2 + \frac{(I_z - I_x)}{I_y}\,p_\omega r_\omega + \frac{C_{\delta_e}}{I_y} V^2 \delta_e \\
&\quad + \frac{t_d}{I_y}\frac{t_p}{V}\,(t_0 + t_1 n_e)\,\delta_{th} \\
\dot{r}_\omega &= \frac{(C_{n\beta}\beta + C_{nr}\,r_\omega)}{I_z} V^2 + \frac{(I_x - I_y)}{I_z}\,p_\omega q_\omega + \frac{C_{\delta_r}}{I_z} V^2 \delta_r \\
\dot{\phi} &= p_\omega + (q_\omega \sin\phi + r_\omega\cos\phi)\tan\theta \\
\dot{\theta} &= q_\omega\cos\phi - r_\omega\sin\phi \\
\dot{\psi} &= \frac{(q_\omega\sin\phi + r_\omega\cos\phi)}{\cos\theta} \\
\dot{n}_e &= t_n n_e^3 + \frac{t_f}{n_e}\,(t_0 + t_1 n_e)\,\delta_{th}
\end{aligned}
\tag{2.42}
$$

where $C_{(\cdot)}$ are the aerodynamic coefficients; $t_{(\cdot)}$ are the engine parameters; w_v, w_l are the vertical and lateral wind disturbance components. The model (2.42) has been obtained on the basis of the following assumptions:

- The expressions of aerodynamic forces and moments have been represented by means of series expansions in the neighbourhood of the steady–state flight condition, then only the main terms are considered.
- The engine model has been simplified by linearising the power with respect to the angular rate behaviour in the neighbourhood of the trim point.
- The second order coupling between the longitudinal and lateral–directional dynamics have been neglected.
- The x–body axis component of the wind has been neglected. In fact, the aircraft behaviour is much more sensitive to the y–body and z–body axis wind components.
- The rudder effect in the equation describing the β dynamics has been neglected. It is worth noting that the designs and the simulations of the NLGA residual generators are robust with respect to this approximation. In fact, the model of the β dynamics will never be used.

Chapter 3

Polynomial Method

In this chapter the PM FDI scheme is presented. The general expression for the residual generator is provided in Section 3.1. An optimisation procedure for the selection of the residual generator parameters is developed in Section 3.2. Finally, a solution for the FDI problem on input–output sensors is proposed in Section 3.3.

3.1 Residual Generation

Let us consider a linear, time-invariant, continuous-time system described by the following input-output equation

$$P(s)\,y(t) = Q_c(s)\,c(t) + Q_d(s)\,d(t) + Q_f(s)\,f(t) \tag{3.1}$$

where $y(t)$ is the m–dimensional output vector, $c(t)$ is the ℓ_c–dimensional known–input vector, $d(t)$ is the ℓ_d–dimensional disturbance vector, $f(t)$ is the ℓ_f–dimensional monitored fault vector. $P(s)$, $Q_c(s)$, $Q_d(s)$, $Q_f(s)$ are polynomial matrices with dimension $(m \times m)$, $(m \times \ell_c)$, $(m \times \ell_d)$, $(m \times \ell_f)$, respectively; $P(s)$ is nonsingular.

Remark 3. *The input–output model (3.1) is obtained from the aircraft linearised state–space model (2.40). Models of type (3.1) are a powerful tool in all fields where the knowledge of the system state does not play a direct role, such as residual generator design, identification, decoupling, output controllability, etc. Algorithms to transform state-space models to equivalent input-output polynomial representations and vice–versa are reported in (Guidorzi 1975).*

A general linear residual generator for the fault detection process of system (3.1) is a filter of type

$$R(s)\,r(t) = S_y(s)\,y(t) + S_c(s)\,c(t) \tag{3.2}$$

that processes the known input–output data and generates the residual $r(t)$, *i.e.* a signal which is "small" (ideally zero) in the fault–free case and is "large" when a fault is acting on the system.

Without loss of generality, $r(t)$ can be assumed to be a scalar signal. In such condition $R(s)$ is a polynomial with degree greater than or equal to the row–degree of $S_c(s)$ and $S_y(s)$, in order to guarantee the physical realisability of the filter.

An important aspect of the design concerns the de–coupling of the disturbance $d(t)$ to produce a correct diagnosis in all operating conditions. If $L(s)$ is a row polynomial vector belonging to $\mathcal{N}_\ell(Q_d(s))$, *i.e.* the left null–space of the matrix $Q_d(s)$, it results

$$L(s)\, Q_d(s)\, d(t) = 0 \tag{3.3}$$

hence pre–multiplying all the terms in (3.1) by $L(s)$, we obtain

$$L(s)\, P(s)\, y(t) - L(s)\, Q_c(s)\, c(t) = L(s)\, Q_f(s)\, f(t) \tag{3.4}$$

Starting from (3.4) with $f(t) = 0$, it is possible to obtain a residual generator of type (3.2) by setting

$$\begin{aligned} S_y(s) &= L(s)\, P(s) \\ S_c(s) &= -L(s)\, Q_c(s) \\ R(s) &= r_1\, s^{n_r} + r_2\, s^{n_r-1} + \ldots + 1 \end{aligned} \tag{3.5}$$

where $n_r \geq n_f$ and n_f is the maximal row–degree of the pair $\{L(s)\, P(s), L(s)\, Q_c(s)\}$. The polynomial $R(s)$ can be arbitrarily selected. The choice of $R(s)$ leads to an asymptotically stable filter when the real parts of the n_r roots are negative. In this way, in absence of fault, relation (3.4) can be rewritten also in the form

$$R(s)\, r(t) = L(s)\, P(s)\, y(t) - L(s)\, Q_c(s)\, c(t) = 0 \tag{3.6}$$

whilst, when a fault is acting on the system, the residual generator is governed by the relation

$$R(s)\, r(t) = L(s)\, Q_f(s)\, f(t) \tag{3.7}$$

and $r(t)$ assumes values that are different from zero if $L(s)$ does not belong to the left null–space of the matrix $Q_f(s)$.

3.1.1 Polynomial Basis Computation

In order to determine all possible residual generators of minimal order, it is necessary to transform model (3.1) into a minimal input-output polynomial representation, that is an equivalent representation with the polynomial matrix $P(s)$ row reduced (Kailath 1980)

$$P(s) = D(s)\, N + E(s) \tag{3.8}$$

where $D(s) = \text{diag}\,\{s^{\nu_1}, s^{\nu_2}, \ldots, s^{\nu_m}\}$ and the highest–row–degree coefficient matrix N is non–singular.

In this condition, the integers ν_i represent the set of the Kronecker output invariants associated to the pair $\{A, C\}$ of every observable realization of $\{P(s), Q(s)\}$ in the state-space. This step can be omitted if the designer is not interested in using minimal order residual generators.

Moreover, it is necessary to compute a minimal basis of $\mathcal{N}_\ell(Q_d(s))$. Under the assumption that matrix $Q_d(s)$ is of full normal rank, *i.e.* $\operatorname{rank} Q_d = \ell_d$, $\mathcal{N}_\ell(Q_d(s))$ has dimension $m - \ell_d$ and a minimal basis of such subspace can be computed as suggested in (Kailath 1980).

It can be noted that in absence of disturbances $\ell_d = 0$ so that $\mathcal{N}_\ell(Q_d(s))$ coincides with the whole vector space. Consequently, a set of residual generators for system (3.1) with $f(t) = 0$ can be expressed as

$$R_{ri}(s)\, r_i(t) = P_{ri}(s)\, y(t) - Q_{c_{ri}}(s)\, c(t) \qquad i = 1, \ldots, m \tag{3.9}$$

where $P_{ri}(s)$ and $Q_{c_{ri}}(s)$ are the i–th rows of matrices $P(s)$ and $Q_c(s)$ respectively, ν_i is the degree of $P_{ri}(s)$ and $R_{ri}(s)$ is an arbitrary polynomial with degree equal to ν_i and with all the roots with negative real part. Since $Q_{c_{ri}}(s)$ cannot show a degree greater than ν_i, the physical realisability of the residual generator is guaranteed.

In general, for $0 < \ell_d < m$ matrix $Q_d(s)$ can be partitioned in the following way

$$Q_d(s) = \begin{bmatrix} Q_{d_1}(s) \\ Q_{d_2}(s) \end{bmatrix} \tag{3.10}$$

where matrices $Q_{d_1}(s)$ and $Q_{d_2}(s)$ have dimensions $\ell_d \times \ell_d$ and $(m - \ell_d) \times \ell_d$ respectively. It can be assumed, without loss of generality, that matrix $Q_{d_1}(s)$ is non singular. In this case it can be easily verified that a basis of $\mathcal{N}_\ell(Q_d(s))$ is given by the polynomial matrix [1]

$$B(s) = \begin{bmatrix} Q_{d_2}(s) \operatorname{adj} Q_{d_1}(s) & -\det Q_{d_1}(s)\, I_{m-\ell_d} \end{bmatrix} \tag{3.11}$$

by assuming $\operatorname{adj} Q_{d_1}(s) = 1$ for $\ell_d = 1$. Note that $B(s)$ has dimension $(m - \ell_d) \times m$. By partitioning $P(s)$ and $Q_c(s)$ as $Q_d(s)$

$$P(s) = \begin{bmatrix} P_1(s) \\ P_2(s) \end{bmatrix} \quad Q_c(s) = \begin{bmatrix} Q_{c_1}(s) \\ Q_{c_2}(s) \end{bmatrix} \tag{3.12}$$

a basis (not necessarily of minimal order) of the residual generator (3.2) for the system (3.1) with $f(t) = 0$ is obtained by replacing in (3.5) the row polynomial vector $L(s)$ with the polynomial matrix $B(s)$, *i.e.*

$$\begin{aligned} S_y(s) &= Q_{d_2}(s) \operatorname{adj} Q_{d_1}(s)\, P_1(s) - \det Q_{d_1}(s)\, P_2(s) \\ S_c(s) &= -Q_{d_2}(s) \operatorname{adj} Q_{d_1}(s)\, Q_{c_1}(s) + \det Q_{d_1}(s)\, Q_{c_2}(s) \\ R(s) &= \operatorname{diag}\{R_1(s), R_2(s), \ldots, R_{m-\ell_d}(s)\} \end{aligned} \tag{3.13}$$

where the degree of the polynomial $R_i(s)$ is n_{f_i}, that is the degree of the i–th row of the matrix $S_y(s)$.

[1] In this work I_n indicates the identity matrix of dimension n, whilst I_n^m indicates the m–th column of I_n.

Remark 4. *By denoting with n_f^* the minimal value of the integers n_{f_i} it is easy to prove that the order n_f^* of a minimal order residual generator for system (3.1) is constrained in the following range*

$$\nu_{min} \leq n_f^* \leq (\ell_d + 1)\,\nu_{max} \tag{3.14}$$

where ν_{min} and ν_{max} are the least and the greatest Kronecker invariant, respectively. The lower bound can be obtained in the no–disturbance case ($\ell_d = 0$) from relation (3.9) by selecting the row of $P(s)$ associated to the least Kronecker invariant. The upper bound can be obtained by taking into account the maximal degree of the polynomials of the matrices. A similar result, obtained with a different approach can be found in (Frisk and Nyberg 2001).

3.1.2 Faults on the Input–Output Sensors

Equation (3.1) considers also the cases of additive faults on the input and output sensors, $f_c(t)$ and $f_o(t)$, respectively. In this situation, only the measurements

$$\begin{aligned} c^*(t) &= c(t) + f_c(t) \\ y^*(t) &= y(t) + f_o(t) \end{aligned} \tag{3.15}$$

are available for the residual generation so that the system (3.1) becomes

$$P(s)\,(y^*(t) - f_o(t)) = Q_c(s)\,(c^*(t) - f_c(t)) + Q_d(s)\,d(t) \tag{3.16}$$

and the residual generator can be written in the following way

$$\begin{aligned} R(s)\,r(t) &= L(s)\,P(s)y^*(t) - L(s)\,Q_c(s)\,c^*(t) \\ &= L(s)\,Q_c(s)\,f_c(t) - L(s)\,P(s)\,f_o(t) \end{aligned} \tag{3.17}$$

Remark 5. *The residual generator described by (3.17) can be seen as an Errors–In–Variables (EIV) model (Van Huffel and Lemmerling 2002) with respect the input and output variable, as the measurements that feed the residual function are affected by additive faults. This description highlights the importance of the residual generator in the form of (3.17)*

3.2 Residual Optimisation

Let us consider the residual generator (3.2) with the positions (3.5) under the assumption that $f(t)$ is a scalar and, consequently, $Q_f(s)$ is a vector

$$\begin{aligned} R(s)\,r(t) &= L(s)\,P(s)\,y(t) - L(s)\,Q_c(s)\,c(t) \\ &= L(s)\,Q_f(s)\,f(t) \end{aligned} \tag{3.18}$$

The diagnostic capabilities of the filter (3.18) strongly depend on the choice of the terms $L(s)$ and $R(s)$. This section proposes a method for the design of these polynomials when $q = m - \ell_d \geq 2$.

The design freedom in the selection of the polynomial row matrix $L(s)$ can be used to optimise the sensitivity properties of $r(t)$ with respect to the fault $f(t)$, for example by maximising the steady–state gain of the transfer function

$$G_f(s) = \frac{L(s)\, Q_f(s)}{R(s)} \tag{3.19}$$

given in (3.18).

In particular, if the row vectors $b_i(s)$ (with $i = 1, \ldots, q$) are the q rows of the basis $B(s)$, $L(s)$ can be expressed as linear combination of these vectors

$$L(s) = \sum_{i=1}^{q} k_i\, b_i(s) \tag{3.20}$$

where k_i are real constants maximising the steady–state gain of the residual generator with respect to the fault, that is

$$\lim_{s\to 0} \frac{1}{R(s)} \left[\sum_{i=1}^{q} k_i\, b_i(s)\right] Q_f(s) = \left[\sum_{i=1}^{q} k_i\, b_i(0)\right] Q_f(0) \tag{3.21}$$

with the constraint

$$\sum_{i=1}^{q} {k_i}^2 = 1 \tag{3.22}$$

In this way, when the fault $f(t)$ is a step–function of magnitude F, the steady–state residual value is

$$\lim_{t\to\infty} r(t) = \lim_{s\to 0} s\, \frac{L(s)\, Q_f(s)}{R(s)}\, \frac{F}{s} = \left[\sum_{i=1}^{q} k_i\, b_i(0)\right] Q_f(0)\, F \tag{3.23}$$

Another design choice regards the location of the roots of the polynomial $R(s)$ in the left–half s–plane, *i.e.* the poles of $G_f(s)$. Since the real coefficients k_i are fixed maximising the steady–state gain there are not design freedom to arbitrarily assign the zeros. In order to solve this problem, in relation (3.20) polynomial coefficients $k_i(s)$ can be considered; in fact, under this assumption, $L(s)$ still belongs to the subspace $\mathcal{N}_\ell(Q_d(s))$. Consequently, in the selection of $L(s)$, there are additional degrees of freedom that can be exploited in order to locate the zeros of $G_f(s)$.

The zeros and poles location influences the transient characteristics (maximum overshoot, delay time, rise time, settling time, etc.) of the residual generator. In many applications, these characteristics must be kept within tolerable or prescribed limits, in order to guarantee good performances of the filter in terms *e.g.* of fault detection times and false alarm rates. This leads to define a poles reference polynomial $U(s)$ and a zeros reference polynomial $H(s)$ whose roots are the poles and the zeros to be assigned, respectively, in order to assure the desired transient characteristics. $R(s)$ and $L(s)$ have to be determined in order to obtain

$$G_f(s) = \frac{H(s)}{U(s)} \tag{3.24}$$

3.2.1 Maximisation of the Steady–State Gain

In this section it is proved the existence and the uniqueness of the solution for the steady–state gain maximisation problem previously formalised. Moreover the analytical computation of this solution is provided.

Since it is conventionally assumed $R(0) = 1$ (see Section 3.1), if the following real vectors are defined

$$k = \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_q \end{bmatrix} \qquad a = B(0)\, Q_f(0) = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_q \end{bmatrix} \tag{3.25}$$

the problem can be recasted as follows.

Problem 1. *Given a, determine k that maximises the steady–state gain, that is, the function $\Phi(k)$ given by the expression*

$$\Phi = k^{\mathrm{T}}\, a = \sum_{i=1}^{q} a_i\, k_i \tag{3.26}$$

under the constraint (3.22).

The constraint (3.22) describes a hypersphere, whilst the function Φ represents a hyperplane. The unknown coefficients k_i must belong to both the hyperplane and the hypersphere. Therefore, the points of tangency between the hypersphere and the hyperplane represents the solutions that maximise or minimise Φ.

Figure 3.1 illustrates the solution of Problem 1 when $q = 2$. In this case the constraint (3.22) is represented by a circle, whilst the expression of the function Φ is a straight line. The unknown coefficients representing the solution must belong to both the circle and the straight line. Since the coefficients a_1 and a_2 are fixed, the position of the straight line is univocally determined by Φ. If Φ increases, the straight line moves to the right, whilst if Φ decreases, the straight line moves to the left. Moreover, if $\Phi = 0$, the straight line passes through origin. Consequently, the point of tangency on the right between the straight line and the circle represents the solution that maximise Φ, whilst the point of tangency on the left represents the solution that minimise Φ. Φ_{max} and Φ_{min}, represented in Figure 3.1, are the maximum and the minimum value of Φ.

In the following, an exact solution for Problem 1 is proposed. Starting from (3.22), k_1 is expressed as a function of $k_2, k_3, \ldots, k_q$ and it is substituted into (3.26)

$$\Phi = a_1 \sqrt{1 - {k_2}^2 - {k_3}^2 - \ \ldots\ - {k_q}^2} + a_2\, k_2 + \ \ldots\ + a_q\, k_q \tag{3.27}$$

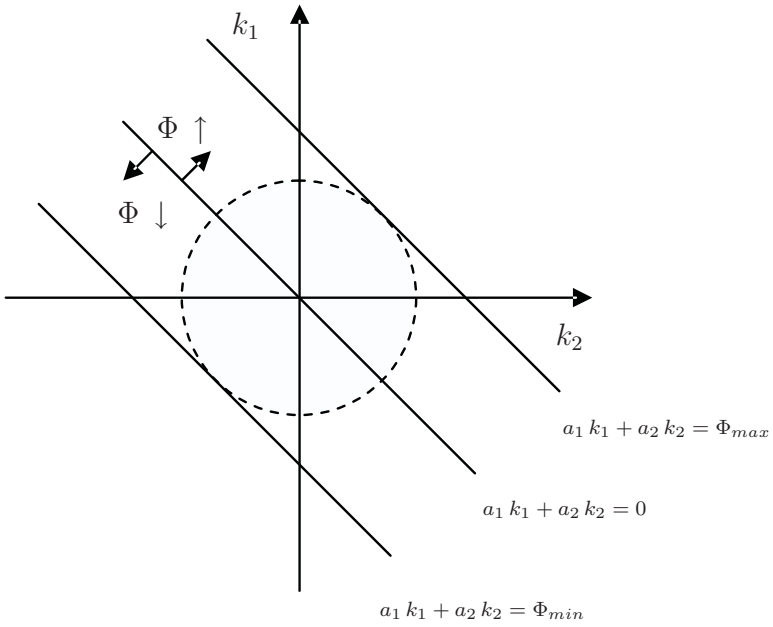


Figure 3.1: Graphical solution of Problem 1 when $q = 2$.

By computing $\nabla\Phi = 0$, *i.e.*

$$\begin{aligned}
\frac{\partial\Phi}{\partial k_2} &= \frac{1}{2}\,a_1\,\frac{-2\,k_2}{\sqrt{1 - {k_2}^2 - {k_3}^2 - \ldots - {k_q}^2}} + a_2 = 0 \\
\frac{\partial\Phi}{\partial k_3} &= \frac{1}{2}\,a_1\,\frac{-2\,k_3}{\sqrt{1 - {k_2}^2 - {k_3}^2 - \ldots - {k_q}^2}} + a_3 = 0 \\
&\vdots \\
\frac{\partial\Phi}{\partial k_q} &= \frac{1}{2}\,a_1\,\frac{-2\,k_q}{\sqrt{1 - {k_2}^2 - {k_3}^2 - \ldots - {k_q}^2}} + a_q = 0
\end{aligned} \tag{3.28}$$

and squaring the expression, after algebraic manipulation

$$\begin{aligned}
{a_2}^2 &= ({a_2}^2 + {a_1}^2){k_2}^2 + {a_2}^2\,{k_3}^2 + \ldots + {a_2}^2\,{k_q}^2 \\
{a_3}^2 &= {a_3}^2\,{k_2}^2 + ({a_3}^2 + {a_1}^2)\,{k_3}^2 + \ldots + {a_3}^2\,{k_q}^2 \\
&\vdots \\
{a_q}^2 &= {a_q}^2\,{k_2}^2 + {a_q}^2\,{k_3}^2 + \ldots + ({a_q}^2 + {a_1}^2)\,{k_q}^2
\end{aligned} \tag{3.29}$$

an expression in the form of $A\,x = b$ is obtained, where

$$A = \begin{bmatrix} ({a_2}^2 + {a_1}^2) & {a_2}^2 & \dots & {a_2}^2 \\ {a_3}^2 & ({a_3}^2 + {a_1}^2) & \dots & {a_3}^2 \\ \vdots & \vdots & \ddots & \vdots \\ {a_q}^2 & {a_q}^2 & \dots & ({a_q}^2 + {a_1}^2) \end{bmatrix}$$

$$x = \begin{bmatrix} {k_2}^2 \\ {k_3}^2 \\ \vdots \\ {k_q}^2 \end{bmatrix} \quad b = \begin{bmatrix} {a_2}^2 \\ {a_3}^2 \\ \vdots \\ {a_q}^2 \end{bmatrix} \tag{3.30}$$

Under the assumption that the constraint (3.22) holds, the vector $\tilde{x}$, representing the squares of the searched Problem 1 solutions, can be expressed as follows

$$\tilde{x} = \begin{bmatrix} 1 - \sum_{i=1}^{q-1} (A^{-1}\,b)_i \\ A^{-1}\,b \end{bmatrix} \tag{3.31}$$

where $(A^{-1}\,b)_i$ is the i–th element of the vector $A^{-1}\,b$.

Let us indicate Ω the set of the vectors k whose elements are the square roots of the elements of $\tilde{x}$. As every element can be taken both with signs '+' and '-', such vectors are 2^q. Therefore, the solution $\tilde{k}$ of Problem 1 can be reformulated as

$$\tilde{k} = \arg \max_{k \in \Omega} \Phi \tag{3.32}$$

In the following it is proposed a Matlab® implementation of an algorithm to find the searched solution $\tilde{k}$ among the 2^q belonging to Ω.

```
function ktilde = fun(x2,a)
q=length(a)
PHImax=0
for l=0:(2^q-1)
  %%% generates x
  for h=1:q
      if bitget(l,h)==0
          x(h)=sqrt(x2(h))
      else
          x(h)=-sqrt(x2(h))
      end
  end
  %%% maximises PHI
  PHI=x*a
  if PHI>PHImax
```

```
        PHImax=PHI
        ktilde=x
    end
end
```

Remark 6. *The matrix A can be expressed as $A = E + {a_1}^2 I_{q-1}$, where E is a matrix with equal columns. If $a_1 \neq 0$, this assumption guarantees the existence of A^{-1}, and consequently the existence and the uniqueness of the solution $A^{-1} b$. Obviously, if $a_1 = 0$ and $a_j \neq 0$, it is sufficient to express k_j as function of the remaining variables and reapply the same procedure.*

Remark 7. *The same solution can be found by maximising the function $|\Phi|$. In fact due to the symmetry properties of the function Φ:*

- $\Phi(k) = \Phi_{max} \Leftrightarrow \Phi(-k) = \Phi_{min}$
- $\Phi_{max} = -\Phi_{min}$

the maximisation of $|\Phi|$ admits two solutions corresponding to the maximum and the minimum of the function Φ.

Remark 8. *Problem 1 could have been solved also in a numerical way,* i.e. *by searching k that maximises Φ on the surface of the q–dimensional hypersphere. However, the computational cost of this numerical solution can be a drawback when q is big.*

3.2.2 Poles and Zeros Assignment

Section 3.2.1 has shown how to maximise the steady–state gain of the transfer function $G_f(s)$ trough a suitable choice of the real vector k (*i.e.* $k = \tilde{k}$). The design of the filter (3.18) has been completed here by introducing a method for assigning both the poles and the zeros of $G_f(s)$.

As it has been said in Section 3.1, $R(s)$ can be arbitrarily selected among the polynomials with degree greater than or equal to n_f (realisability condition) and with all the roots in the left–half s–plane (stability condition). Moreover, it is conventionally assumed $R(0) = 1$. Consequently, if the poles reference polynomial $U(s)$ satisfies these conditions, the poles are assigned imposing $R(s) = U(s)$.

Let us consider the zeros assignment problem. According to Section 3.2.1 the q–dimensional polynomial vector $a(s) = B(s)\, Q_f(s)$ is defined. The i–th element of this vector is a known polynomial of a certain degree called n_{a_i}. Note that if n_a is defined as follows

$$n_a = \max_{i=1,\ldots,q} n_{a_i} \tag{3.33}$$

the i–th element of $a(s)$ can be always written as a polynomial of degree n_a

$$a_i(s) = \sum_{j=0}^{n_a} a_i^j \, s^j \tag{3.34}$$

by imposing that $a_i^j = 0$ when $j > n_{a_i}$. It is defined also the q–dimensional polynomial vector $k(s)$ whose i–th element has the form

$$k_i(s) = \sum_{j=0}^{n_k} k_i^j \, s^j \tag{3.35}$$

Since $L(s)$ can be expressed as linear combination of the rows of $B(s)$ with polynomial coefficients $k_i(s)$, *i.e.* $L(s) = k^{\mathrm{T}}(s)\, B(s)$, the degree n_k and the $q \times (n_k + 1)$ coefficients k_i^j are degrees of freedom that can be exploited by the designer in order to obtain desired roots for $L(s)\, Q_f(s) = k^{\mathrm{T}}(s)\, a(s)$. However, in order to maximise the steady–state gain, as shown in Section 3.2.1, the following constraint have to be satisfied

$$k(0) = \tilde{k} = \begin{bmatrix} \tilde{k_1} \\ \tilde{k_2} \\ \vdots \\ \tilde{k_q} \end{bmatrix} \iff \quad k_i^0 = \tilde{k_i} \quad i = 1, \ldots, q \tag{3.36}$$

It is worth noting that due to the constraint (3.36) the zeros reference polynomial, defined as follows

$$H(s) = \sum_{j=0}^{n_h} h^j \, s^j \tag{3.37}$$

must satisfy the condition $H(0) = \tilde{k}^{\mathrm{T}}\, a(0)$. Obviously this assumption does not provide any restriction on the roots assignable.

Under the previous hypotheses the zeros assignment problem can be formulated in the following way.

Problem 2. *Given $a(s)$ and $H(s)$, find the degree n_k and the coefficients k_i^j, under the constraint (3.36), in order to obtain $k^{\mathrm{T}}(s)\, a(s) = H(s)$.*

By multiplying (3.35) and (3.34), it results

$$k^{\mathrm{T}}(s)\, a(s) = \sum_{i=1}^{q} \sum_{j=0}^{n_k+n_a} \left(\sum_{\alpha+\beta=j} k_i^\alpha \, a_i^\beta \right) s^j = \sum_{j=0}^{n_k+n_a} e^j \, s^j \tag{3.38}$$

where

$$e^j = \sum_{i=1}^{q} \sum_{\alpha+\beta=j} k_i^\alpha \, a_i^\beta \tag{3.39}$$

In (3.38) and (3.39) it is assumed that $k_i^\alpha = 0$ when $\alpha > n_k$ and $a_i^\beta = 0$ when $\beta > n_a$. Note that the coefficients $e^1, \ldots, e^{n_k+n_a}$ depend on the freedom design $k_i^1, \ldots, k_i^{n_k}$. On the other hand, e^0 is fixed as the coefficients k_i^0 are assigned by the constraint (3.36).

Let us suppose $n_h \leq n_k + n_a$. By imposing $k^{\mathrm{T}}(s)\, a(s) = H(s)$, from (3.39) and (3.37), the following expressions are computed

$$\sum_{i=1}^{q} \sum_{\alpha+\beta=j} k_i^{\alpha}\, a_i^{\beta} = h^j \quad j = 0, \ldots, n_k + n_a \tag{3.40}$$

where it is supposed $h^j = 0$ when $j = n_h + 1, \ldots, n_k + n_a$. The relations of (3.36) and (3.40) represent a linear system, with $n_k + n_a$ equations and $q \times n_k$ unknowns, that can be expressed in the classical form $A\, x = b$, where

$$A = \begin{bmatrix} a_1^0 & \ldots & a_q^0 & 0 & \ldots & 0 & & 0 & \ldots & 0 \\ \vdots & \ddots & \vdots & a_1^0 & \ldots & a_q^0 & & & & \\ a_1^{n_a} & \ldots & a_q^{n_a} & \vdots & \ddots & \vdots & & & & \\ 0 & \ldots & 0 & a_1^{n_a} & \ldots & a_q^{n_a} & & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & 0 & \ldots & 0 & & & & \\ & & & & & & \ddots & & & \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & & 0 & \ldots & 0 \\ & & & & & & & a_1^0 & \ldots & a_q^0 \\ & & & & & & & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & 0 & \ldots & 0 & & a_1^{n_a} & \ldots & a_q^{n_a} \end{bmatrix}$$

$$x = \begin{bmatrix} k_1^1 \\ \vdots \\ k_q^1 \\ k_1^2 \\ \vdots \\ k_q^2 \\ \vdots \\ \vdots \\ k_1^{n_k} \\ \vdots \\ k_q^{n_k} \end{bmatrix} \qquad b = \begin{bmatrix} h^1 - \sum_{i=1}^{q} k_i^0 a_i^1 \\ \vdots \\ h^{n_a} - \sum_{i=1}^{q} k_i^0 a_i^{n_a} \\ h^{n_a+1} \\ \vdots \\ h^{n_a+n_k} \end{bmatrix} \tag{3.41}$$

The degree n_k of the polynomials $k_i(s)$ has to be chosen in order to obtain a solvable system (*i.e.* $\operatorname{rank} A = \operatorname{rank} [A\ b]$). An automatic procedure to properly choose n_k and consequently to solve Problem 2 is showed in Figure 3.2. To understand the proposed procedure the following points should be considered:

- The choice of n_k must guarantee that the hypotheses $n_h \leq n_k + n_a$ is satisfied.
- When $q \geq 2$, the difference between the number of unknown terms and the number of equations, *i.e.* $(q-1) \times n_k - n_a - 1$, is greater than zero if n_k is selected sufficiently large.
- Even if the system admits solutions, the inverse of the matrix A may not exist. In such case there are infinite solutions and the one associated to the pseudo-inverse of A, *i.e.* $A^+ b$ can be considered.

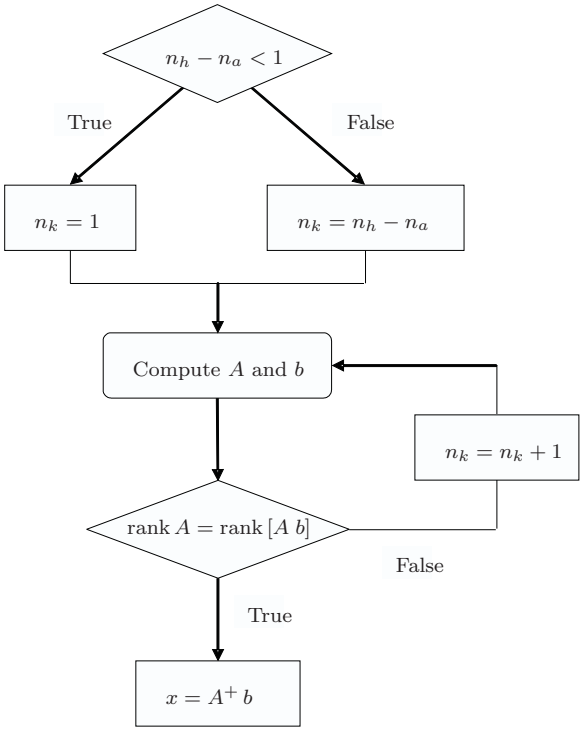


Figure 3.2: Automatic procedure to solve Problem 2.

Remark 9. *The use of a polynomial vector $k(s)$ instead of a real vector k has the drawback of increasing the complexity of the residual generator. Many FDI applications require* $n_h = 0$, i.e.

$$G_f(s) = \frac{H(0)}{U(s)} \tag{3.42}$$

In such cases it is not needed to find $k(s)$ such that $k^{\mathrm{T}}(s)\, a(s) = H(0)$ but it is easier considering $k = \tilde{k}$ and imposing

$$R(s) = \frac{\tilde{k}^{\mathrm{T}}\, a(s)\, U(s)}{H(0)} \tag{3.43}$$

Obviously, due to the realisability condition, it must be $\deg\{U(s)\} \geq n_f - \deg\{\tilde{k}^{\mathrm{T}}\, a(s)\}$. *Moreover the method cannot be applied if* $\tilde{k}^{\mathrm{T}}\, a(s)$ *admits one or more roots in the right–half s–plane, in fact the residual generator would result unstable. In such cases an approximate solution can be developed (Beghelli* et al. *2007*a*).*

Remark 10. *The problems (and the relative solution) discussed in this section in the continuos–time domain, can be easily extended to the discrete–time domain (Simani and Benini 2007). The main difference between the two approaches can be identified when the polynomial k method is needed. In fact, in order to maximise the steady–state gain, in the continuous–time case it is required* $k(s) = \tilde{k}$ *when* $s = 0$*, whilst in the discrete–time case it is required* $k(z) = \tilde{k}$ *when* $z = 1$*. Obviously this is a consequence of the fact that the final value theorem changes if the continuous–time domain or the discrete–time domain is considered.*

Remark 11. *Section 3.2 is focused on the design of residual generators on the basis of a given reference function with disturbance decoupling and fault sensitivity maximisation properties. The pole location influences the transient dynamics of the designed residual filters, while the steady–state properties depend on the PM residual design, as it maximises the residual steady–state values with respect to step faults affecting input and output sensors. The poles of the residual functions could be optimised with respect to both fault and disturbance terms, as shown* e.g. *in (Bonfè* et al. *2004).*

3.3 FDI on Input–Output Sensors

This section addresses the problem of the design of a bank of residual generators for the isolation of faults affecting the input and output sensors. The design is performed by using the disturbance de–coupling method suggested in Section 3.1. In the following it is assumed that $m > \ell_d + 1$.

3.3.1 Bank for Input Sensors FDI

To univocally isolate a fault concerning one of the input sensors, under the hypothesis that the remaining input sensors and all output sensors are fault–free, a bank of residual generator filters is used, according to Figure 3.3. The number of these generators is equal to the number ℓ_c of system control inputs, and the i–th device ($i = 1, \ldots, \ell_c$) is driven by all but the i–th input and all the outputs of the system. In this case, a fault on the i–th input sensor affects all but the i–th residual generator.

With reference to Figure 3.3, $c^{*i}(t)$ represents the $(\ell_c - 1)$–dimensional vector obtained by deleting from $c^*(t)$ the i–th component, with

$$c^*(t) = c(t) + f_{c_i}(t) \tag{3.44}$$

and

$$f_{c_i}(t) = \left[\begin{array}{ccccccc} 0 & \ldots & 0 & h_{c_i}(t) & 0 & \ldots & 0 \end{array}\right]^{\mathrm{T}} \tag{3.45}$$

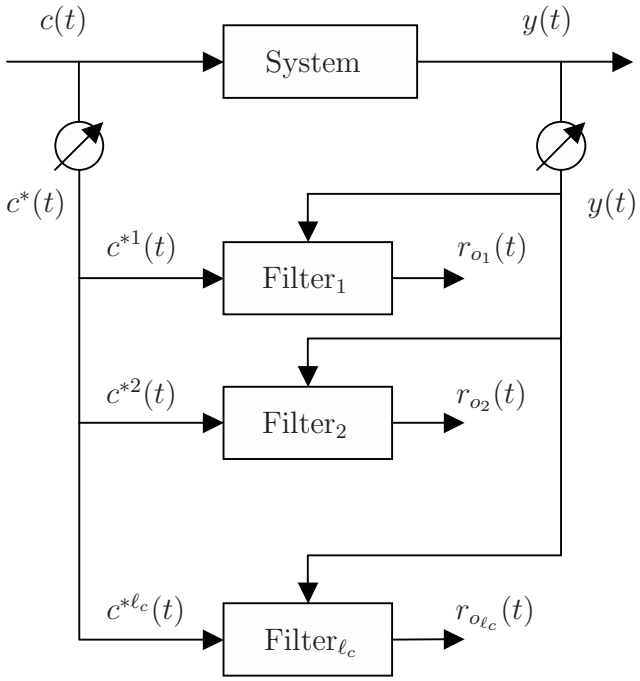


Figure 3.3: Bank of filters for fault isolation on the input sensors.

When the fault on the i–th input sensor $h_{c_i}(t)$ is considered, the system (3.1) can be rewritten as follows

$$P(s)\,y(t) = Q_c(s)\,c(t) + Q_d(s)\,d(t) + q_{c_i}(s)\,h_{c_i}(t) \tag{3.46}$$

where $q_{c_i}(s)$ represents the i–th column of the matrix $Q_c(s)$.

Hence, by multiplying relation (3.46) by the matrix $L_{c_i}(s)$, where $L_{c_i}(s)$ is a row vector belonging to the basis for the left null space of the matrix $\left[Q_d(s)\,|\,q_{c_i}(s)\right]$, and $Q_c^i(s)$ is the matrix obtained by deleting from $Q_c(s)$ the i–th column, the equation of the i–th filter becomes

$$R_{c_i}(s)\,r_{c_i}(t) = L_{c_i}(s)\,P(s)\,y(t) - L_{c_i}(s)\,Q_c^i(s)\,c^{*i}(t) = 0 \tag{3.47}$$

whilst, for the j–th filter, with $j \neq i$, it results

$$\begin{aligned} R_{c_j}(s)\,r_{c_j}(t) &= L_{c_j}(s)\,P(s)\,y(t) - L_{c_j}(s)\,Q_c^j(s)\,c^{*j}(t) \\ &= L_{c_j}(s)\,q_{c_i}(s)\,h_{c_i}(t) \end{aligned} \tag{3.48}$$

$R_{c_i}(s)$ and $R_{c_j}(s)$ are arbitrary polynomials with all the roots with negative real part.

3.3.2 Bank for Output Sensors FDI

In order to univocally isolate a fault concerning one of the output sensors, under the hypotheses that all the input sensors and the remaining output sensors are fault–free, a

bank of residual generator filters is used, according to Figure 3.4. The number of these generators is equal to the number m of system outputs, and the i–th device ($i = 1, \ldots, m$) is driven by all but the i–th output and all the inputs of the system. In this case, a fault on the i–th output sensor affects all but the i–th residual generator.

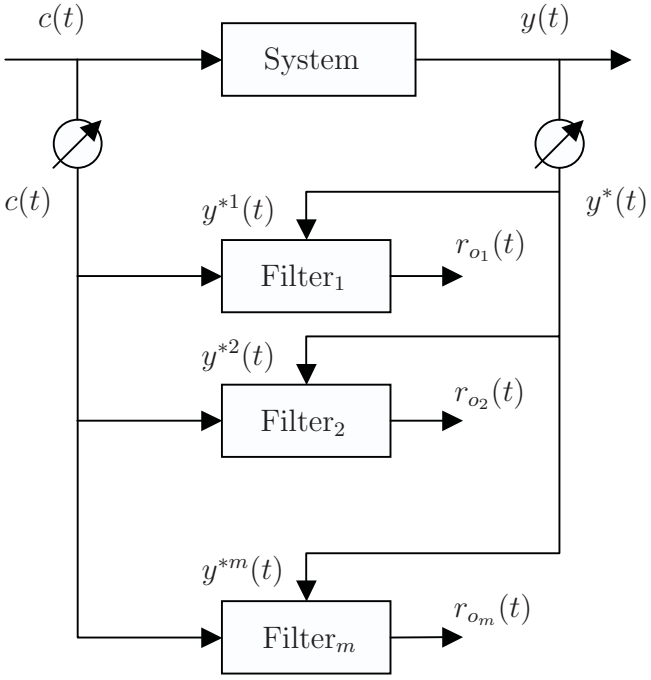


Figure 3.4: Bank of filters for fault isolation on the output sensors.

With reference to Figure 3.4, $y^{*i}(t)$ represents the $(m-1)$–dimensional vector obtained by deleting from $y^*(t)$ the i–th component, with

$$y^*(t) = y(t) + f_{o_i}(t) \tag{3.49}$$

and

$$f_{o_i}(t) = \left[\begin{array}{ccccccc} 0 & \ldots & 0 & h_{o_i}(t) & 0 & \ldots & 0 \end{array} \right]^{\mathrm{T}} \tag{3.50}$$

When the fault on the i–th output sensor $h_{o_i}(t)$ is considered, the system (3.1) can be rewritten as follows

$$P(s)\, y(t) = Q_c(s)\, c(t) + Q_d(s)\, d(t) - p_i(s)\, h_{o_i}(t) \tag{3.51}$$

where $p_i(s)$ represents the i–th column of the matrix $P(s)$.

Hence, by multiplying relation (3.51) by the matrix $L_{o_i}(s)$, where $L_{o_i}(s)$ is a row vector belonging to the basis for the left null space of the matrix $\left[Q_d(s) \,|\, p_i(s)\right]$, and denoting $P^i(s)$ the matrix obtained by deleting from $P(s)$ the i–th column, the equation of the i–th filter becomes

$$R_{o_i}(s)\, r_{o_i}(t) = L_{o_i}(s)\, P^i(s)\, y^{*i}(t) - L_{o_i}(s)\, Q_c(s)\, c(t) = 0 \tag{3.52}$$

whilst, for the j–th filter, with $j \neq i$, it results

$$\begin{aligned} R_{o_j}(s)\, r_{o_j}(t) &= L_{o_j}(s)\, P^j(s)\, y^{*j}(t) - L_{o_j}(s)\, Q_c(s)\, c(t) \\ &= -L_{o_j}(s)\, p_i(s)\, h_{o_i}(t) \end{aligned} \tag{3.53}$$

$R_{o_i}(s)$ and $R_{o_j}(s)$ are arbitrary polynomials whose roots have negative real part.

3.3.3 Fault Signature

In order to summarise the FDI capabilities of the presented schemes, Table 3.1 shows the fault signatures in case of a single fault in each input and output sensor.

Table 3.1: Fault signatures.

Residual / Fault	f_{c_1}	f_{c_2}	$\dots$	$f_{c_{\ell_c}}$	f_{o_1}	f_{o_2}	$\dots$	f_{o_m}
r_{c_1}	0	1	$\dots$	1	1	1	$\dots$	1
r_{c_2}	1	0	$\dots$	1	1	1	$\dots$	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$r_{c_{\ell_c}}$	1	1	$\dots$	0	1	1	$\dots$	1
r_{o_1}	1	1	$\dots$	1	0	1	$\dots$	1
r_{o_2}	1	1	$\dots$	1	1	0	$\dots$	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
r_{o_m}	1	1	$\dots$	1	1	1	$\dots$	0

The residuals which are affected by input and output faults are marked with the presence of '1' in the correspondent table entry, while an entry '0' means that the input or output fault does not affect the correspondent residual. All the elements out of the main diagonal on Table 3.1 are '1' when both the following conditions hold:

- For $i = 1, \dots, \ell_c$, the column vectors of the matrix $Q_c^i(s)$ and the column vectors of the matrix $P(s)$ are not orthogonal with the row vector $L_{c_i}(s)$.
- For $j = 1, \dots, m$, the column vectors of the matrix $P^j(s)$ and the column vectors of the matrix $Q_c(s)$ are not orthogonal with the row vector $L_{o_j}(s)$.

When not all the elements out of the main diagonal of the Table 3.1 are '1's, the fault isolation is still feasible if the columns of the fault signature table are all different from each other.

Remark 12. *From the comparison between the filter (3.18) and the generic filter of the input bank given by (3.48), the following associations can be made*

$$R(s) = R_{c_j}(s) \quad L(s) = L_{c_j}(s) \quad Q_f(s) = q_{c_i}(s) \quad f(t) = h_{c_i}(t) \tag{3.54}$$

whilst, from the comparison with the generic filter of the output bank given by (3.53), it results

$$R(s) = R_{o_j}(s) \quad L(s) = L_{o_j}(s) \quad Q_f(s) = p_i(s) \quad f(t) = h_{o_i}(t) \tag{3.55}$$

Hence if $q = m - l_d - 1 \geq 2$*, the optimisation method shown in Section 3.2 and can be exploited for the design of the* j*–th filter of the input or output bank. In particular, the parameters of this filter can be properly chosen in order to optimise its performances when a fault is acting on the* i*–th input or output sensor.*

Chapter 4

Nonlinear Geometric Approach

In this chapter the NLGA–based FDI schems are developed. The classical NLGA technique is proposed in Section 4.1. A procedure to improve the robustness of the NLGA scheme is presented in Section 4.2. Finally, the NLGA–AF and the NLGA–PF techniques are developed in Sections 4.3 and 4.4, respectively.

4.1 NLGA FDI Scheme

The NLGA approach to nonlinear FDI problem is suggested in (De Persis and Isidori 2000) and formally developed in (De Persis and Isidori 2001). It consists in finding, by means of a coordinate change in the state space and in the output space, an observable subsystem which, if possible, is affected by the fault and not affected by disturbance. In this way, necessary and sufficient conditions for the FDI problem to be solvable are given. Finally, a residual generator can be designed on the basis of the model of the observable subsystem. This technique was applied for the first time to a Vertical Take–Off and Landing (VTOL) aircraft with reference to a reduced–order model (De Persis *et al.* 2001).

4.1.1 Coordinate Transformation

The approach consider a nonlinear system model in the form

$$\begin{aligned} \dot{x} &= n(x) + g(x)\, c + \ell(x)\, f + p(x)\, d \\ y &= h(x) \end{aligned} \tag{4.1}$$

in which $x \in \mathcal{X}$ (an open subset of $\Re^n$) is the state vector, $c(t) \in \Re^{\ell_c}$ is the control input vector, $f(t) \in \Re$ is the fault, $d(t) \in \Re^{\ell_d}$ the disturbance vector (embedding also the faults which have to be decoupled) and $y \in \Re^m$ the output vector. $n(x)$, $\ell(x)$, the columns of $g(x)$ and $p(x)$ are smooth vector fields; and $h(x)$ is a smooth map.

Therefore, if P represents the distribution spanned by the column of $p(x)$, the NLGA method can be stated as follows (De Persis and Isidori 2001):

1. Determine the minimal conditioned invariant distribution containing P (denoted with Σ_*^P).

2. By using $(\Sigma_*^P)^\perp$, *i.e.* the maximal conditioned invariant codistribution contained in $P^\perp$, determine the largest observability codistribution contained in $P^\perp$, denoted with Ω^*.

3. If $\ell(x) \notin \Omega^*$ continue to the next step, otherwise the fault is not detectable.

4. If the condition of the previous step is satisfied, it can be found a surjection Ψ_1 and a function Φ_1 fulfilling $\Omega^* \cap \text{span}\{dh\} = \text{span}\{d(\Psi_1 \circ h)\}$ and $\Omega^* = \text{span}\{d(\Phi_1)\}$, respectively. The functions $\Psi(y)$ and $\Phi(x)$ defined as

$$\Psi(y) = \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \end{pmatrix} = \begin{pmatrix} \Psi_1(y) \\ H_2\, y \end{pmatrix} \qquad \Phi(x) = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{pmatrix} = \begin{pmatrix} \Phi_1(x) \\ H_2\, h(x) \\ \Phi_3(x) \end{pmatrix} \tag{4.2}$$

   are (local) diffeomorphisms, where H_2 is a selection matrix (*i.e.* a matrix in which any row has all 0 entries but one, which is equal to 1), $\Phi_1(x)$ represents the measured part of the state which is affected by f and not affected by d and $\Phi_3(x)$ represents the not measured part of the state which is affected by f and by d.

In the new (local) coordinate defined previously, the system (4.1) is described by the relations in the form

$$\begin{aligned}
\dot{\bar{x}}_1 &= n_1(\bar{x}_1, \bar{x}_2) + g_1(\bar{x}_1, \bar{x}_2)\, c + \ell_1(\bar{x}_1, \bar{x}_2, \bar{x}_3)\, f \\
\dot{\bar{x}}_2 &= n_2(\bar{x}_1, \bar{x}_2, \bar{x}_3) + g_2(\bar{x}_1, \bar{x}_2, \bar{x}_3)\, c + \ell_2(\bar{x}_1, \bar{x}_2, \bar{x}_3)\, f + p_2(\bar{x}_1, \bar{x}_2, \bar{x}_3)\, d \\
\dot{\bar{x}}_3 &= n_3(\bar{x}_1, \bar{x}_2, \bar{x}_3) + g_3(\bar{x}_1, \bar{x}_2, \bar{x}_3)\, c + \ell_3(\bar{x}_1, \bar{x}_2, \bar{x}_3)\, f + p_3(\bar{x}_1, \bar{x}_2, \bar{x}_3)\, d \\
\bar{y}_1 &= h(\bar{x}_1) \\
\bar{y}_2 &= \bar{x}_2
\end{aligned} \tag{4.3}$$

with $\ell_1(\bar{x}_1, \bar{x}_2, \bar{x}_3)$ not identically zero. Denoting $\bar{x}_2$ with $\bar{y}_2$ and considering it as an independent input, it can be singled out the $\bar{x}_1$–subsystem

$$\begin{aligned}
\dot{\bar{x}}_1 &= n_1(\bar{x}_1, \bar{y}_2) + g_1(\bar{x}_1, \bar{y}_2)\, c + \ell_1(\bar{x}_1, \bar{y}_2, \bar{x}_3)\, f \\
\bar{y}_1 &= h(\bar{x}_1)
\end{aligned} \tag{4.4}$$

which is affected by the single fault f and decoupled from the disturbance vector. This subsystem has been exploited for the design of the residual generator for the FDI of the fault f.

4.1.2 Residual Generators Design

As already described in Section 2.3.2, the proposed NLGA FDI scheme is designed on the basis of model structure of the input affine type as expressed by (4.1). For this reason the aircraft simulation model has been simplified and the nonlinear model (2.42) has been considered for the NLGA design. From the comparison between (4.1) and (2.42), the following positions hold

$$\begin{aligned} x = y &= \begin{bmatrix} V & \alpha & \beta & p_\omega & q_\omega & r_\omega & \phi & \theta & \psi & n_e \end{bmatrix}^{\mathrm{T}} \\ c &= \begin{bmatrix} \delta_e & \delta_a & \delta_r & \delta_{th} \end{bmatrix}^{\mathrm{T}} \end{aligned} \tag{4.5}$$

hence $h(x) = I_{10}$. Moreover

$$n(x) = \begin{bmatrix} -\frac{\left(C_{D0}+C_{D\alpha}\alpha+C_{D\alpha_2}\alpha^2\right)}{m}V^2 + g\left(\sin\alpha\cos\theta\cos\phi - \cos\alpha\sin\theta\right) \\ -\frac{(C_{L0}+C_{L\alpha}\alpha)}{m}V + \frac{g}{V}\left(\cos\alpha\cos\theta\cos\phi + \sin\alpha\sin\theta\right) + q_\omega \\ \frac{\left(C_{D0}+C_{D\alpha}\alpha+C_{D\alpha_2}\alpha^2\right)\sin\beta + C_{Y\,\beta}\beta\cos\beta}{m}V + g\frac{\cos\theta\sin\phi}{V} + p_\omega\sin\alpha - r_\omega\cos\alpha \\ \frac{\left(C_{l\beta}\beta+C_{lp}\,p_\omega\right)}{I_x}V^2 + \frac{(I_y-I_z)}{I_x}\,q_\omega r_\omega \\ \frac{(C_{m0}+C_{m\alpha}\alpha+C_{mq}\,q_\omega)}{I_y}V^2 + \frac{(I_z-I_x)}{I_y}\,p_\omega r_\omega \\ \frac{\left(C_{n\beta}\beta+C_{nr}\,r_\omega\right)}{I_z}V^2 + \frac{(I_x-I_y)}{I_z}\,p_\omega q_\omega \\ p_\omega + \left(q_\omega\sin\phi + r_\omega\cos\phi\right)\tan\theta \\ q_\omega\cos\phi - r_\omega\sin\phi \\ \frac{(q_\omega\sin\phi + r_\omega\cos\phi)}{\cos\theta} \\ t_n n_e^3 \end{bmatrix} \tag{4.6}$$

and

$$\begin{aligned} g(x) &= \begin{bmatrix} g_1(x) & g_2(x) & g_3(x) & g_4(x) \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 0 & \frac{\cos\alpha}{m}\frac{t_p}{V}\left(t_0 + t_1 n_e\right) \\ 0 & 0 & 0 & -\frac{\sin\alpha}{m}\frac{t_p}{V^2}\left(t_0 + t_1 n_e\right) \\ 0 & 0 & 0 & -\frac{\cos\alpha\sin\beta}{m}\frac{t_p}{V^2}\left(t_0 + t_1 n_e\right) \\ 0 & \frac{C_{\delta_a}}{I_x}V^2 & 0 & 0 \\ \frac{C_{\delta_e}}{I_y}V^2 & 0 & 0 & \frac{t_d}{I_y}\frac{t_p}{V}\left(t_0 + t_1 n_e\right) \\ 0 & 0 & \frac{C_{\delta_r}}{I_z}V^2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{t_f}{n_e}\left(t_0 + t_1 n_e\right) \end{bmatrix} \end{aligned} \tag{4.7}$$

The distribution matrix $p_d(x)$ of the vertical and lateral wind disturbance components w_v and w_l is also defined

$$p_d(x) = \begin{bmatrix} \sin\alpha & \frac{\cos\alpha}{V} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{V} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^{\mathrm{T}} \tag{4.8}$$

In the following the residual generators for detecting the faults affecting the aircraft input sensors are obtained.

Elevator Residual Generator Design

To decouple the elevator residual generator from the wind and faults on aileron, rudder and throttle, the distribution P is defined as

$$p(x) = \begin{bmatrix} p_d(x) & g_2(x) & g_3(x) & g_4(x) \end{bmatrix} \tag{4.9}$$

Hence, the closure of P is given by $\bar{P} = [P\ I_{10}^{10}]$. Now, by recalling that $\mathrm{Ker}\{dh\} = \emptyset$, it follows that $\Sigma_*^P = \bar{P}$. Hence $(\Sigma_*^P)^\perp = (\bar{P})^\perp$ is given by

$$(\bar{P})^\perp = \begin{bmatrix} \cos\alpha & -V\sin\alpha & 0 & 0 & -\frac{I_y}{mt_d} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{4.10}$$

By observing that $\mathrm{span}\{dh\} = I_{10}$, it follows that $\Omega^* = (\Sigma_*^P)^\perp = (\bar{P})^\perp$, hence $(\Omega^*)^\perp = \bar{P}$. Because $\ell(x) = g_1(x) \notin (\Omega^*)^\perp$, the fault is detectable. The change of output coordinates gives

$$\Psi_1(x) = \bar{x}_1 = \begin{bmatrix} V\cos\alpha - \frac{I_y}{mt_d} q_\omega \\ \phi \\ \theta \\ \psi \end{bmatrix} \qquad H_2 x = \bar{x}_2 = \begin{bmatrix} V \\ \alpha \\ \beta \\ p_\omega \\ r_\omega \\ n_e \end{bmatrix} \tag{4.11}$$

Note that only the first component of the vector $\bar{x}_1$, *i.e.* $\bar{x}_{11}$, is directly affected by the fault. In fact the other variables are not fed by the inputs. In order to design the residual generator it is necessary to compute

$$\begin{aligned} \dot{\bar{x}}_{11} &= \dot{V}\cos\alpha - V\dot{\alpha}\sin\alpha - \frac{I_y}{mt_d}\dot{q}_\omega \\ &= \frac{V^2}{m}\left[-\left(C_{D0} + C_{D\alpha}\alpha + C_{D\alpha_2}\alpha^2\right)\cos\alpha\right] + \frac{V^2}{m}\left(C_{L0} + C_{L\alpha}\alpha\right)\sin\alpha \\ &- g\sin\theta - Vq_\omega\sin\alpha - \frac{(C_{m0} + C_{m\alpha}\alpha + C_{mq}q_\omega)}{mt_d}V^2 - \frac{(I_z - I_x)}{mt_d}p_\omega r_\omega \\ &- \frac{C_{\delta_e}}{mt_d}V^2\delta_e \end{aligned} \tag{4.12}$$

Hence, with $k_{\delta_e} > 0$, the elevator residual generator r_{δ_e} is given by

$$\begin{aligned}\dot{\xi}_1 &= \frac{V^2}{m}\left[-\left(C_{D0} + C_{D\alpha}\alpha + C_{D\alpha_2}\alpha^2\right)\cos\alpha\right] + \frac{V^2}{m}\left(C_{L0} + C_{L\alpha}\alpha\right)\sin\alpha \\ &\quad - g\sin\theta - Vq_\omega\sin\alpha - \frac{(C_{m0} + C_{m\alpha}\alpha + C_{mq}q_\omega)}{mt_d}V^2 - \frac{(I_z - I_x)}{mt_d}p_\omega r_\omega \\ &\quad - \frac{C_{\delta_e}}{mt_d}V^2\delta_e + k_{\delta_e}\left[\left(V\cos\alpha - \frac{I_y}{mt_d}q_\omega\right) - \xi_1\right] \\ r_{\delta_e} &= \left(V\cos\alpha - \frac{I_y}{mt_d}q_\omega\right) - \xi_1\end{aligned} \tag{4.13}$$

Aileron Residual Generator Design

To decouple the aileron residual generator from the wind and faults on elevator, rudder and throttle, the distribution P is defined as

$$p(x) = \begin{bmatrix} p_d(x) & g_1(x) & g_3(x) & g_4(x) \end{bmatrix} \tag{4.14}$$

Hence, the closure of P is given by $\bar{P} = [P\ I_{10}^{10}]$. Now, by recalling that $\mathrm{Ker}\{dh\} = \emptyset$, it follows that $\Sigma_*^P = \bar{P}$. Hence $(\Sigma_*^P)^\perp = (\bar{P})^\perp$ is given by

$$(\bar{P})^\perp = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{4.15}$$

By observing that $\mathrm{span}\{dh\} = I_{10}$, it follows that $\Omega^* = (\Sigma_*^P)^\perp = (\bar{P})^\perp$, hence $(\Omega^*)^\perp = \bar{P}$. Because $\ell(x) = g_2(x) \notin (\Omega^*)^\perp$, the fault is detectable. The change of output coordinates gives

$$\Psi_1(x) = \bar{x}_1 = \begin{bmatrix} p_\omega \\ \phi \\ \theta \\ \psi \end{bmatrix} \qquad H_2 x = \bar{x}_2 = \begin{bmatrix} V \\ \alpha \\ \beta \\ q_\omega \\ r_\omega \\ n_e \end{bmatrix} \tag{4.16}$$

Note that only $\bar{x}_{11}$ is directly affected by the fault. In fact the other variables are not fed by the inputs. In order to design the residual generator it is necessary to compute

$$\dot{\bar{x}}_{11} = \dot{p}_\omega = \frac{(C_{l\beta}\beta + C_{lp}\,p_\omega)}{I_x}V^2 + \frac{(I_y - I_z)}{I_x}\,q_\omega r_\omega + \frac{C_{\delta_a}}{I_x}V^2\delta_a \tag{4.17}$$

Hence, with $k_{\delta_a} > 0$, the aileron residual generator r_{δ_a} is given by

$$\begin{aligned}\dot{\xi}_2 &= \frac{(C_{l\beta}\beta + C_{lp}p_\omega)}{I_x}V^2 + \frac{(I_y - I_z)}{I_x}q_\omega r_\omega + \frac{C_{\delta_a}}{I_x}V^2\delta_a + k_{\delta_a}\left(p_\omega - \xi_2\right) \\ r_{\delta_a} &= p_\omega - \xi_2\end{aligned} \tag{4.18}$$

Rudder Residual Generator Design

To decouple the rudder residual generator from the wind and faults on elevator, aileron and throttle, the distribution P is defined as

$$p(x) = \begin{bmatrix} p_d(x) & g_1(x) & g_2(x) & g_4(x) \end{bmatrix} \tag{4.19}$$

Hence, the closure of P is given by $\bar{P} = [P\ I_{10}^{10}]$. Now, by recalling that $\mathrm{Ker}\{dh\} = \emptyset$, it follows that $\Sigma_*^P = \bar{P}$. Hence $(\Sigma_*^P)^\perp = (\bar{P})^\perp$ is given by

$$(\bar{P})^\perp = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{4.20}$$

By observing that $\mathrm{span}\{dh\} = I_{10}$, it follows that $\Omega^* = (\Sigma_*^P)^\perp = (\bar{P})^\perp$, hence $(\Omega^*)^\perp = \bar{P}$. Because $\ell(x) = g_3(x) \notin (\Omega^*)^\perp$, the fault is detectable. The change of output coordinates gives

$$\Psi_1(x) = \bar{x}_1 = \begin{bmatrix} r_\omega \\ \phi \\ \theta \\ \psi \end{bmatrix} \qquad H_2 x = \bar{x}_2 = \begin{bmatrix} V \\ \alpha \\ \beta \\ p_\omega \\ q_\omega \\ n_e \end{bmatrix} \tag{4.21}$$

Note that only $\bar{x}_{11}$ is directly affected by the fault. In fact the other variables are not fed by the inputs. In order to design the residual generator it is necessary to compute

$$\dot{\bar{x}}_{11} = \dot{r}_\omega = \frac{(C_{n\beta}\beta + C_{nr}\, r_\omega)}{I_z} V^2 + \frac{(I_x - I_y)}{I_z}\, p_\omega q_\omega + \frac{C_{\delta_r}}{I_z} V^2 \delta_r \tag{4.22}$$

Hence, with $k_{\delta_r} > 0$, the rudder residual generator r_{δ_r} is given by

$$\begin{aligned} \dot{\xi}_3 &= \frac{(C_{n\beta}\beta + C_{nr} r_\omega)}{I_z} V^2 + \frac{(I_x - I_y)}{I_z} p_\omega q_\omega + \frac{C_{\delta_r}}{I_z} V^2\, \delta_r + k_{\delta_r}\, (r_\omega - \xi_3) \\ r_{\delta_r} &= r_\omega - \xi_3 \end{aligned} \tag{4.23}$$

Throttle Residual Generator Design

To decouple the throttle residual generator from the wind and faults on elevator, aileron and rudder, the distribution P is defined as

$$p(x) = \begin{bmatrix} p_d(x) & g_1(x) & g_2(x) & g_3(x) \end{bmatrix} \tag{4.24}$$

Since P is an involutive distribution, it results $\bar{P} = P$. Now, by recalling that $\mathrm{Ker}\{dh\} = \emptyset$, it follows that $\Sigma_*^P = \bar{P}$. Hence $(\Sigma_*^P)^\perp = (\bar{P})^\perp$ is given by

$$(\bar{P})^\perp = \begin{bmatrix} \cos\alpha & -V\sin\alpha & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{4.25}$$

By observing that $\mathrm{span}\{dh\} = I_{10}$, it follows that $\Omega^* = (\Sigma_*^P)^\perp = (\bar{P})^\perp$, hence $(\Omega^*)^\perp = \bar{P}$. Because $\ell(x) = g_4(x) \notin (\Omega^*)^\perp$, the fault is detectable. The change of output coordinates gives

$$\Psi_1(x) = \bar{x}_1 = \begin{bmatrix} V\cos\alpha \\ \phi \\ \theta \\ \psi \\ n_e \end{bmatrix} \qquad H_2 x = \bar{x}_2 = \begin{bmatrix} V\sin\alpha \\ \beta \\ p_\omega \\ q_\omega \\ r_\omega \end{bmatrix} \tag{4.26}$$

Note that both $\bar{x}_{15}$ and $\bar{x}_{11}$ are affected by the fault, leading to two throttle residual generators:

- To design the residual generator related to $\bar{x}_{15}$, it is necessary to compute

$$\dot{\bar{x}}_{15} = \dot{n}_e = t_n n_e^3 + \frac{t_f}{n_e}\left(t_0 + t_1 n_e\right)\delta_{th} \tag{4.27}$$

  Hence, with $k_{\delta_{th}} > 0$, the rudder residual generator $r_{\delta_{th}}$ related to $\bar{x}_{15}$ is given by

$$\begin{aligned} \dot{\xi}_4 &= t_n n_e^3 + \frac{t_f}{n_e}\left(t_0 + t_1 n_e\right)\delta_{th} + k_{\delta_{th}}\left(n_e - \xi_4\right) \\ r_{\delta_{th}} &= n_e - \xi_4 \end{aligned} \tag{4.28}$$

- To design the residual generator related to $\bar{x}_{11}$, it is necessary to compute

$$\begin{aligned} \dot{\bar{x}}_{11} &= \dot{V}\cos\alpha - V\dot{\alpha}\sin\alpha \\ &= -\frac{(C_{d0} + C_{d\alpha}\alpha + C_{d\alpha 2}\alpha^2)}{m} V^2\cos\alpha + V^2\sin\alpha\frac{(C_{L0} + C_{L\alpha}\alpha)}{m} \\ &- g\sin\theta - V q_\omega\sin\alpha + \frac{t_p}{mV}\left(t_0 + t_1 n_e\right)\delta_{th} \end{aligned} \tag{4.29}$$

  Hence, with $k'_{\delta_{th}} > 0$, the rudder residual generator $r'_{\delta_{th}}$ related to $\bar{x}_{11}$ is given by

$$\begin{aligned} \dot{\xi}'_4 &= -\frac{(C_{d0} + C_{d\alpha}\alpha + C_{d\alpha 2}\alpha^2)}{m} V^2\cos\alpha + V^2\sin\alpha\frac{(C_{L0} + C_{L\alpha}\alpha)}{m} \\ &- g\sin\theta - V q_\omega\sin\alpha + \frac{t_p}{mV}\left(t_0 + t_1 n_e\right)\delta_{th} + k'_{\delta_{th}}(V\cos\alpha - \xi'_4) \\ r'_{\delta_{th}} &= (V\cos\alpha - \xi'_4) \end{aligned} \tag{4.30}$$

The residual generator $r_{\delta_{th}}$ is characterised by a fewer number of parameters with respect to $r'_{\delta_{th}}$. Hence the choice of $r_{\delta_{th}}$ is preferable to cope with robustness requirements. However it also possible to use jointly the two residual generator.

Remark 13. *Each residual generator is affected by a single input sensor fault and is decoupled from the wind components and the faults affecting the remaining input sensors. In this way the tuning of the residual generator gains k_{δ_e}, k_{δ_a}, k_{δ_r} and $k_{\delta_{th}}$ can be carried out independently. Finally, by a straightforward analysis, the positive sign of each gain is a necessary and sufficient condition for the asymptotic stability of the designed residual generators.*

A procedure optimising the trade–off between the fault sensitivity and the robustness to the modelling errors and disturbances of the generic residual generator is proposed in the next section.

4.2 NLGA Robustness Improvements

The proposed NLGA FDI based scheme consists of two design steps:

1. The structural decoupling of critical disturbances and critical modelling errors can be obtained as described in Section 4.1.

2. The nonlinear residual generators robustness is improved by minimising the effects of both non critical disturbances and modelling errors, whilst maximising the fault effects on the residual signals.

4.2.1 The $\bar{x}_{11}$–subsystem

In order to apply the robustness improvement procedure presented in this section, the considered framework is restricted to suitable scalar components of the $\bar{x}_1$–subsystem (4.4). In particular, the vectors $\bar{x}_1$ and $\bar{y}_1$ are decomposed as follows

$$\bar{x}_1 = \begin{bmatrix} \bar{x}_{11} \\ \bar{x}_{1\mathrm{c}} \end{bmatrix} \qquad \bar{y}_1 = \begin{bmatrix} \bar{y}_{11} \\ \bar{y}_{1\mathrm{c}} \end{bmatrix} \tag{4.31}$$

where $\bar{x}_{11} \in \Re$, $\bar{y}_{11} \in \Re$ and correspondingly it follows

$$n_1(\cdot) = \begin{bmatrix} n_{11}(\cdot) \\ n_{1\mathrm{c}}(\cdot) \end{bmatrix} \qquad g_1(\cdot) = \begin{bmatrix} g_{11}(\cdot) \\ g_{1\mathrm{c}}(\cdot) \end{bmatrix} \qquad \ell_1(\cdot) = \begin{bmatrix} \ell_{11}(\cdot) \\ \ell_{1\mathrm{c}}(\cdot) \end{bmatrix} \tag{4.32}$$

Let us consider the following conditions

$$\bar{y}_{11} = h_{11}(\bar{x}_{11}) \qquad \bar{y}_{1\mathrm{c}} = h_{1\mathrm{c}}(\bar{x}_{1\mathrm{c}}) \qquad \ell_{11}(\cdot) \neq 0 \tag{4.33}$$

where $h_{11}(\cdot)$ is a smooth map and $h_{1c}(\cdot)$ is an invertible smooth map. It is important to highlight that if the constraints (4.33) are satisfied, the decomposition (4.31)–(4.32) can always be applied to obtain the following $\bar{x}_{11}$–subsystem

$$\begin{aligned} \dot{\bar{x}}_{11} &= n_{11}(\bar{x}_{11}, \bar{y}_{1c}, \bar{y}_2) + g_{11}(\bar{x}_{11}, \bar{y}_{1c}, \bar{y}_2)c + \ell_{11}(\bar{x}_{11}, \bar{y}_{1c}, \bar{y}_2, \bar{x}_3)f \\ \bar{y}_{11} &= h_{11}(\bar{x}_{11}) \end{aligned} \tag{4.34}$$

As can be seen in Section 4.1.2, the conditions (4.33) are satisfied for the considered aircraft application, hence, from now on, the scalar $\bar{x}_{11}$–subsystem (4.34) is referred to in place of the $\bar{x}_1$–subsystem (4.4).

It can be noted that the tuning of the residual generator gains, in the framework of the $\bar{x}_{11}$–subsystem (4.34), cannot be properly carried out. In fact the critical disturbances are structurally decoupled but the non critical ones are not considered. For this reason, to achieve robustness of the residual generators, the tuning of the gains is performed by embedding the description of the non critical disturbances in the $\bar{x}_{11}$–subsystem as follows

$$\begin{aligned} \dot{\bar{x}}_{11} &= n_{11}(\bar{x}_{11}, \bar{y}_{1c}, \bar{y}_2) + g_{11}(\bar{x}_{11}, \bar{y}_{1c}, \bar{y}_2)c + \ell_{11}(\bar{x}_{11}, \bar{y}_{1c}, \bar{y}_2, \bar{x}_3)f \\ &\quad + e(\bar{x}_{11}, \bar{y}_{1c}, \bar{y}_2, \bar{x}_3)\zeta \\ \bar{y}_{11} &= \bar{x}_{11} + \nu \end{aligned} \tag{4.35}$$

where, to simplify the treatment without loss of generality (accordingly to the considered aircraft application), the state variable $\bar{x}_{11}$ is supposed to be directly measured. Moreover, the variable $\nu \in \Re$ is the measurement noise on $\bar{x}_{11}$. Finally, the variable $\zeta \in \Re$ and the related scalar field $e(\cdot)$ represent the non critical effects which have not been considered in the simplified aircraft model (2.42) used for the NLGA scheme.

4.2.2 Filter Form and Observer Form

The following system, which is referred to as filter form, represents a generic scalar residual generator (based on the subsystem (4.35)) to which the residual generators designed in Section 4.1.2 belong as a particular case

$$\begin{aligned} \dot{\xi}_f &= n_{11}(\bar{y}_{11}, \bar{y}_{1c}, \bar{y}_2) + g_{11}(\bar{y}_{11}, \bar{y}_{1c}, \bar{y}_2)c + k_f\,(\bar{y}_{11} - \xi_f) \\ r_f &= \bar{y}_{11} - \xi_f \end{aligned} \tag{4.36}$$

where the gain k_f has to be tuned in order to minimise the effects of the disturbances ζ and ν whilst maximise the effects of the fault f on the residual r_f.

The quantification both of the disturbances and of the fault effects on the residual can be obtained by defining the estimation error

$$\tilde{x}_f = \bar{x}_{11} - \xi_f \tag{4.37}$$

which allows to write the following equivalent residual model

$$\begin{aligned} \dot{\tilde{x}}_f &= n_{11}(\bar{x}_{11}, \bar{y}_{1c}, \bar{y}_2) - n_{11}(\bar{y}_{11}, \bar{y}_{1c}, \bar{y}_2) + g_{11}(\bar{x}_{11}, \bar{y}_{1c}, \bar{y}_2)c - g_{11}(\bar{y}_{11}, \bar{y}_{1c}, \bar{y}_2)c \\ &\quad + \ell_{11}(\bar{x}_{11}, \bar{y}_{1c}, \bar{y}_2, \bar{x}_3)f + e(\bar{x}_{11}, \bar{y}_{1c}, \bar{y}_2, \bar{x}_3)\zeta - k_f\tilde{x}_f - k_f\nu \\ r_f &= \tilde{x}_f + \nu \end{aligned} \tag{4.38}$$

In order to apply the effective mixed $\mathcal{H}_-/\mathcal{H}_\infty$ approach (Chen and Patton 1999, Hou and Patton 1996) to tune k_f, the system (4.38) has to be linearised in the neighbourhood of a stationary flight condition, as suggested in (Amato *et al.* 2006) with reference to the $\mathcal{H}_\infty$ optimisation of nonlinear unknown input observers. It is worth observing that the considered aircraft application is characterised by small excursions of the state, input and output variables with respect to their trim values $\bar{x}_{10}$, $\bar{x}_{30}$, c_0, $\bar{y}_{10}$ and $\bar{y}_{20}$, hence the robustness of the nonlinear residual generator is achieved. The linearisation of (4.38) is the following

$$\begin{aligned} \dot{\tilde{x}}_f &= -k_f \tilde{x}_f - k_f \nu + mf + \breve{q}\breve{\zeta} \\ r_f &= \tilde{x}_f + \nu \end{aligned} \tag{4.39}$$

where

$$\begin{aligned} a' &= \left.\frac{\partial n_{11}(\cdot)}{\partial \bar{x}_{11}}\right|_{(\bar{x}_{10},\bar{y}_{20})} \qquad b = g_{11}(\cdot)|_{(\bar{x}_{10},\bar{y}_{20})} \\ m &= \ell_{11}(\cdot)|_{(\bar{x}_{10},\bar{y}_{20},\bar{x}_{30})} \qquad q = e(\cdot)|_{(\bar{x}_{10},\bar{y}_{20},\bar{x}_{30})} \end{aligned} \tag{4.40}$$

and

$$\breve{q}\breve{\zeta} = q\zeta - a'\nu \tag{4.41}$$

Now, it is important to note that in place of the residual generators in the filter form (4.36), the following observer form of the residual generators can be used

$$\begin{aligned} \dot{\xi}_o &= n_{11}(\xi_o, \bar{y}_{1c}, \bar{y}_2) + g_{11}(\xi_o, \bar{y}_{1c}, \bar{y}_2)c + k_o\left(\bar{y}_{11} - \xi_o\right) \\ r_o &= \bar{y}_{11} - \xi_o \end{aligned} \tag{4.42}$$

For the same reasons previously described, the estimation error $\tilde{x}_o$ is introduced

$$\tilde{x}_o = \bar{x}_{11} - \xi_o \tag{4.43}$$

hence

$$\begin{aligned} \dot{\tilde{x}}_o &= n_{11}(\bar{x}_{11}, \bar{y}_{1c}, \bar{y}_2) - n_{11}(\xi_o, \bar{y}_{1c}, \bar{y}_2) + g_{11}(\bar{x}_{11}, \bar{y}_{1c}, \bar{y}_2)c - g_{11}(\xi_o, \bar{y}_{1c}, \bar{y}_2)c \\ &\quad + \ell_{11}(\bar{x}_{11}, \bar{y}_{1c}, \bar{y}_2, \bar{x}_3)f + e(\bar{x}_{11}, \bar{y}_{1c}, \bar{y}_2, \bar{x}_3)\zeta - k_o\tilde{x}_o - k_o\nu \\ r_o &= \tilde{x}_o + \nu \end{aligned} \tag{4.44}$$

and the linearisation of (4.44) is

$$\begin{aligned} \dot{\tilde{x}}_o &= \left(a' - k_o\right)\tilde{x}_o - k_o\nu + mf + q\zeta \\ r_o &= \tilde{x}_o + \nu \end{aligned} \tag{4.45}$$

Both the linearised models (4.39) and (4.45) of the residual generators in the filter form and observer form respectively can be represented by the following general form

$$\begin{aligned} \dot{\tilde{x}} &= \left(a - k\right)\tilde{x} + \left(E_1 - kE_2\right)\varepsilon + mf \\ r &= \tilde{x} + E_2\varepsilon \end{aligned} \tag{4.46}$$

with $E_1 = [e_{11}\ 0]$ as well as the following positions

$$\begin{array}{lccccccc} \text{general form} & \tilde{x} & \varepsilon & r & a & k & e_{11} & E_2 \\ \text{filter form} & \tilde{x}_f & [\breve{\zeta}\ \nu]^{\mathrm{T}} & r_f & 0 & k_f & \breve{q} & [0\ 1] \\ \text{observer form} & \tilde{x}_0 & [\zeta\ \nu]^{\mathrm{T}} & r_0 & a' & k_0 & q & [0\ 1] \end{array} \tag{4.47}$$

On the basis of (4.46) and (4.47), the mixed $\mathcal{H}_-/\mathcal{H}_\infty$ (Chen and Patton 1999, Hou and Patton 1996) procedure is developed for the robustness improvement of the residual generators both in the filter and observer form. Since the considered NLGA residual generators are scalar, the $\mathcal{H}_-/\mathcal{H}_\infty$ procedure leads to a new analytical solution.

4.2.3 Mixed $\mathcal{H}_-/\mathcal{H}_\infty$ Optimisation

Let us define the norms $\mathcal{H}_\infty$ and $\mathcal{H}_-$ of a stable transfer function G as

$$\|G\|_\infty = \sup_{\omega \geq 0} \bar{\sigma}\left[G\left(j\,\omega\right)\right] \qquad \|G\|_- = \underline{\sigma}\left[G\left(j\,0\right)\right] \tag{4.48}$$

where $\bar{\sigma}$ and $\underline{\sigma}$ represents the maximum and the minimum singular value, respectively. The problem of the trade–off between disturbances robustness and fault sensitivity is stated as follows.

Problem 3. *Given two scalars $\beta > 0$ and $\gamma > 0$, find the set $\mathcal{K}$ defined as*

$$\mathcal{K} = \left\{k \in \Re :\ (a-k) < 0,\ \|G_{r\varepsilon}\|_\infty < \gamma,\ \|G_{rf}\|_- > \beta\right\} \tag{4.49}$$

where

$$G_{r\,\varepsilon}(s) = \left(s - a + k\right)^{-1}\ \left(E_1 - k\,E_2\right) + E_2 \tag{4.50}$$

and

$$G_{rf}(s) = \left(s - a + k\right)^{-1}\ m \tag{4.51}$$

In order to obtain the analytical solution of Problem 3, the following propositions are given.

Proposition 1. $\forall k \in \Re,\ (a-k) < 0,$ *then*

$$\|G_{r\varepsilon}\|_\infty^2 = \max\left\{1,\ \frac{(e_{11}^2 + a^2)}{(k-a)^2}\right\} \tag{4.52}$$

and

$$\sup_{\{k \in \Re:\,(a-k)<0\}} \|G_{r\varepsilon}\|_\infty = +\infty \tag{4.53}$$

Proof. From the definition (4.50)

$$G_{r\varepsilon}(s) = \left[\frac{e_{11}}{s-a+k} \quad \frac{s-a}{s-a+k}\right] \tag{4.54}$$

hence it is possible to write

$$\left\{\bar{\sigma}\left[G_{r\varepsilon}(j\,\omega)\right]\right\}^2 = \frac{e_{11}^2}{(k-a)^2+\omega^2} + \frac{a^2+\omega^2}{(k-a)^2+\omega^2} = \frac{(e_{11}^2+a^2)+\omega^2}{(k-a)^2+\omega^2} \tag{4.55}$$

so that it follows

$$\|G_{r\varepsilon}\|_\infty^2 = \sup_{\xi\geq 0} \frac{(e_{11}^2+a^2)+\xi}{(k-a)^2+\xi} \tag{4.56}$$

From the last expression, it is straightforward to obtain (4.52) and (4.53). □

Proposition 2. *The set*

$$\mathcal{K}_\gamma = \{k \in \Re : (a-k) < 0,\ \|G_{r\varepsilon}\|_\infty < \gamma,\ \gamma > 1\} \tag{4.57}$$

is given by

$$k > \underline{k} \quad \textit{with} \quad \underline{k} = a + \frac{\sqrt{e_{11}^2+a^2}}{\gamma} \tag{4.58}$$

Proof. By means of Proposition 1, it is possible to write

$$\frac{(e_{11}^2+a^2)}{(k-a)^2} < \gamma^2 \tag{4.59}$$

which holds for

$$k > a + \frac{\sqrt{e_{11}^2+a^2}}{\gamma} \tag{4.60}$$

□

Proposition 3. *If* $\gamma > 1$*, then* $\left\{\|G_{rf}\|_- : \|G_{r\varepsilon}\|_\infty < \gamma\right\}$ *is given by*

$$0 < \|G_{rf}\|_- < \beta_{\max}(\gamma) \quad \textit{with} \quad \beta_{\max}(\gamma) = \frac{m\gamma}{\sqrt{e_{11}^2+a^2}} \tag{4.61}$$

Proof. From the definition (4.51), it results $G_{rf}(s) = m/(s-a+k)$ and assuming, without loss of generality, that $m > 0$, it follows $\|G_{rf}\|_- = m/(k-a)$. By imposing $\|G_{rf}\|_- > \beta$ with $\beta > 0$, the constraint $k < a + (m/\beta)$ has to hold. Then, by recalling the result of Proposition 2, the maximum feasible value of β fulfilling the constraint $\|G_{r\varepsilon}\|_\infty < \gamma$ is given by

$$\underline{k} = a + \frac{m}{\beta_{\max}(\gamma)} \tag{4.62}$$

hence

$$\beta_{\max}(\gamma) = \frac{m}{\underline{k}-a} = \frac{m\gamma}{\sqrt{e_{11}^2+a^2}} \tag{4.63}$$

□

Theorem 1. *Given $\gamma > 1$ and $\beta \in]0, \beta_{\max}(\gamma)[$, the set $\mathcal{K}$ fulfilling the constraints of Problem 3 is given by*

$$\mathcal{K} = \left\{ k \in \mathcal{R} : k \in \left]\underline{k}, \overline{k}\right[, \underline{k} = a + \frac{m}{\beta_{\max}(\gamma)}, \overline{k} = a + \frac{m}{\beta} \right\} \tag{4.64}$$

Proof. The proof of the theorem is not reported, as it is straightforward from Propositions 1, 2 and 3. □

Remark 14. *Let us consider the following performance index to maximise*

$$J = \frac{\|G_{rf}\|_-}{\|G_{r\varepsilon}\|_\infty} \tag{4.65}$$

From (4.52) it follows

$$\|G_{r\varepsilon}\|_\infty = \begin{cases} 1 & k > \left(a + \sqrt{e_{11}^2 + a^2}\right) \\ \dfrac{\sqrt{e_{11}^2 + a^2}}{k - a} & a < k \leq \left(a + \sqrt{e_{11}^2 + a^2}\right) \end{cases} \tag{4.66}$$

hence

$$J = \begin{cases} \dfrac{m}{k - a} & k > \left(a + \sqrt{e_{11}^2 + a^2}\right) \\ \dfrac{m}{\sqrt{e_{11}^2 + a^2}} & a < k \leq \left(a + \sqrt{e_{11}^2 + a^2}\right) \end{cases} \tag{4.67}$$

From (4.67), it can be observed that

$$J = \frac{m}{k - a} < \frac{m}{\sqrt{e_{11}^2 + a^2}}, \qquad k > \left(a + \sqrt{e_{11}^2 + a^2}\right) \tag{4.68}$$

In this way the maximum value of the performance index J is

$$J_{\max} = \frac{m}{\sqrt{e_{11}^2 + a^2}} \qquad \forall k \in \mathcal{K}_J = \left\{ k \in \Re : a < k \leq \left(a + \sqrt{e_{11}^2 + a^2}\right) \right\} \tag{4.69}$$

The method proposed in this work guarantees the maximum value of the performance index J as well as the constraints $\|G_{r\varepsilon}\|_\infty < \gamma$ and $\|G_{rf}\|_- > \beta$, if $\beta \geq m/\sqrt{e_{11}^2 + a^2}$. In fact from $\beta \geq m/\sqrt{e_{11}^2 + a^2}$ it follows

$$\|G_{rf}\|_- = \frac{m}{k - a} > \beta \geq \frac{m}{\sqrt{e_{11}^2 + a^2}} \tag{4.70}$$

hence $k < \left(a + \sqrt{e_{11}^2 + a^2}\right)$. Finally, from (4.61) it is always possible to find a β such that

$$\frac{m}{\sqrt{e_{11}^2 + a^2}} \leq \beta \leq \beta_{max}(\gamma) \qquad \forall \gamma > 1 \tag{4.71}$$

On the basis of Theorem 1, the residual generator gain k can be designed by means of the following procedure:

1. Choose $\gamma > 1$ to obtain a desired level of disturbance attenuation.
2. Compute $\beta_{\max}(\gamma)$ and choose $\beta \in \,]0, \beta_{\max}(\gamma)[$ to obtain a desired level of fault sensitivity.
3. Choose $k \in \left]\underline{k}, \overline{k}\right[$, with $\underline{k} = a + m/\beta_{\max}(\gamma)$ and $\overline{k} = a + m/\beta$.
4. Apply the chosen gain k to the k_f of (4.36) or to the k_o of (4.42) if the NLGA residual generator is in the filter form or in the observer form respectively.

4.3 NLGA–AF FDI Scheme

The NLGA–AF FDI scheme belongs to the NLGA framework, where the coordinate transformation detailed in Section 4.1 is the starting point to design a set of adaptive filters in order to detect an additive fault acting on a single input sensor and to estimate the magnitude of the fault.

The information brought by the fault size estimation can be very useful for off–line maintenance purposes and for on–line reconfiguration of the automatic flight control system. It is worth observing that the basic NLGA scheme based on residual signals cannot provide fault size estimation.

Nonlinear geometric approaches with estimation of the fault size can be found also in (Kaboré *et al.* 2000, Kaboré and Wang 2001), in which the fault estimation method relies on the successive derivatives of input/output signals. A drawback of this strategy is a high sensitivity to measurement noise.

4.3.1 Adaptive Filtering Algorithm

In the following an adaptive nonlinear filter for the $\bar{x}_1$–subsystem, providing fault size estimation, is developed. Moreover, the asymptotic convergence of the estimate to the actual fault size is formally proven.

Remark 15. *The NLGA–AF FDI scheme can be applied only if the fault detectability condition presented in Section 4.1.1 holds and the following new constraints are satisfied:*

- *The $\bar{x}_1$–subsystem is independent from the $\bar{x}_3$ state components.*
- *The fault is a step function of the time, hence the parameter f is a constant to be estimated.*
- *There exists a proper scalar component $\bar{x}_{1s}$ of the state vector $\bar{x}_1$ such that the corresponding scalar component of the output vector is $\bar{y}_{1s} = \bar{x}_{1s}$ and the following relation holds*

$$\dot{\bar{y}}_{1s}(t) = M_1(t) \cdot f + M_2(t) \tag{4.72}$$

where $M_1(t) \neq 0, \forall t \geq 0$. Moreover $M_1(t)$ and $M_2(t)$ can be computed for each time instant, since they are functions just of input and output measurements. The relation (4.72) describes the general form of the system under diagnosis.

Problem 4. *The design of an adaptive filter is required, with reference to the system model (4.72), in order to perform an estimation $\hat{f}(t)$, which asymptotically converges to the magnitude of the fault f.*

The proposed adaptive filter that solves the FDI Problem 4 is based on the least–squares algorithm with forgetting factor and described by the following adaptation law

$$\begin{aligned} \dot{P} &= \beta\, P - \frac{1}{N^2} P^2 \breve{M}_1^2 \qquad & P\,(0) &= P_0 > 0 \\ \dot{\hat{f}} &= P\, \epsilon\, \breve{M}_1 & \hat{f}\,(0) &= 0 \end{aligned} \tag{4.73}$$

with the following equations representing the output estimation and the corresponding normalised estimation error

$$\begin{aligned} \hat{\breve{y}}_{1s} &= \breve{M}_1\, \hat{f} + \breve{M}_2 + \lambda\, \breve{\bar{y}}_{1s} \\ \epsilon &= \frac{1}{N^2} \left(\bar{y}_{1s} - \hat{\breve{y}}_{1s} \right) \end{aligned} \tag{4.74}$$

where all the involved variables of the adaptive filter are scalar. In particular, $\lambda > 0$ is a parameter related to the bandwidth of the filter, $\beta \geq 0$ is the forgetting factor and $N^2 = 1 + \breve{M}_1^2$ is the normalisation factor of the least–squares algorithm.

Moreover, the proposed adaptive filter adopts the signals $\breve{M}_1$, $\breve{M}_2$, $\breve{\bar{y}}_{1s}$ which are obtained by means of a low–pass filtering of the signals M_1, M_2, $\bar{y}_{1s}$ as follows

$$\begin{aligned} \dot{\breve{M}}_1 &= -\lambda\, \breve{M}_1 + M_1 \qquad & \breve{M}_1(0) &= 0 \\ \dot{\breve{M}}_2 &= -\lambda\, \breve{M}_2 + M_2 & \breve{M}_2(0) &= 0 \\ \dot{\breve{\bar{y}}}_{1s} &= -\lambda\, \breve{\bar{y}}_{1s} + \bar{y}_{1s} & \breve{\bar{y}}_{1s}(0) &= 0 \end{aligned} \tag{4.75}$$

Proposition 4. *The considered adaptive filter is described by (4.73)–(4.75). The asymptotic relation between the normalised output estimation error $\epsilon(t)$ and the fault estimation error $f - \hat{f}(t)$ is the following*

$$\lim_{t \to \infty} \epsilon(t) = \lim_{t \to \infty} \frac{\breve{M}_1(t)}{N^2(t)} \left(f - \hat{f}(t) \right) \tag{4.76}$$

Proof. The following auxiliary system is defined in the form

$$\begin{aligned} \dot{y}'_1 &= -\lambda\, y'_1 + \dot{\bar{y}}_{1s} \qquad & y'_1(0) &= 0 \\ \dot{y}'_2 &= -\lambda\, y'_2 + \lambda\, \bar{y}_{1s} & y'_2(0) &= 0 \\ y' &= y'_1 + y'_2 \end{aligned} \tag{4.77}$$

By means of simple computations, it follows

$$\begin{aligned} y'(t) &= \int_0^t e^{-\lambda(t-\tau)}\dot{\bar{y}}_{1s}(\tau)d\tau + \int_0^t e^{-\lambda(t-\tau)}\lambda\,\bar{y}_{1s}(\tau)d\tau \\ &= \int_0^t e^{-\lambda(t-\tau)}\big(M_1(\tau)f + M_2(\tau)\big)d\tau + \lambda\,\breve{\bar{y}}_{1s} \\ &= \breve{M}_1(t)f + \breve{M}_2(t) + \lambda\,\breve{\bar{y}}_{1s}(t) \end{aligned} \tag{4.78}$$

Let us consider the following function

$$V = \frac{1}{2}\,(y' - \bar{y}_{1s})^2 \tag{4.79}$$

which is trivially positive definite and radially unbounded. Moreover, its first time derivative is

$$\begin{aligned} \dot{V} &= (y' - \bar{y}_{1s})(\dot{y}'_1 + \dot{y}'_2 - \dot{\bar{y}}_{1s}) \\ &= (y' - \bar{y}_{1s})(-\lambda\,y'_1 - \lambda\,y'_2 + \lambda\,\bar{y}_{1s}) \\ &= -\lambda\,(y' - \bar{y}_{1s})^2 \end{aligned} \tag{4.80}$$

Since $\dot{V}$ is trivially negative definite $\forall\ y' \neq \bar{y}_{1s}$, V is a Lyapunov function so that $y'(t)$ globally asymptotically tends to the output function $\bar{y}_{1s}(t)$ and from (4.78) the following relation holds

$$\lim_{t\to\infty} \bar{y}_{1s}(t) = \breve{M}_1(t)f + \breve{M}_2(t) + \lambda\,\breve{\bar{y}}_{1s}(t) \tag{4.81}$$

From (4.74) and from the expression (4.81), the asymptotic behaviour of the normalised output estimation error $\epsilon(t)$ can be straightforwardly obtained as follows

$$\begin{aligned} \lim_{t\to\infty} \epsilon(t) &= \lim_{t\to\infty} \frac{1}{N^2(t)} \left(\bar{y}_{1s}(t) - \breve{M}_1(t)\hat{f}(t) - \breve{M}_2(t) - \lambda\,\breve{\bar{y}}_{1s}(t)\right) \\ &= \lim_{t\to\infty} \frac{1}{N^2(t)} \left(\breve{M}_1(t)f - \breve{M}_1(t)\hat{f}(t)\right) \end{aligned} \tag{4.82}$$

□

Theorem 2. *The adaptive filter described by (4.73)–(4.75) represents a solution to the FDI Problem 4, so that $\hat{f}(t)$ provides an asymptotically convergent estimation of the magnitude of the step fault f.*

Proof. Let us consider the following function

$$W = \frac{1}{2}\left(\hat{f} - f\right)^2 \tag{4.83}$$

which is trivially positive definite and radially unbounded. Moreover, its first time derivative is

$$\dot{W} = \left(\hat{f} - f\right)\left(P\,\epsilon\,\breve{M}_1 - 0\right) \tag{4.84}$$

It is worth noting that the smoothness property of the involved functions allows to apply the asymptotic approximation (4.76) to the expression (4.84). In fact, $\exists\ t_\star > 0$ so that the sign of $\dot{W}(t)$, $\forall\ t \geq t_\star$ is not affected by the asymptotic approximation (4.76). Hence it follows

$$\dot{W}(t) = -P(t)\frac{\breve{M}_1^2(t)}{N^2(t)}\left(\hat{f}(t) - f\right)^2 \qquad \forall\ t \geq t_\star \tag{4.85}$$

which is negative definite $\forall\ \hat{f} \neq f$. In fact, $\breve{M}_1(t)$ is a low–pass filtering of the signal $M_1(t)$ which is a smooth function and always not null by hypothesis in the Remark 15. Moreover $N^2(t) = 1 + \breve{M}_1^2(t) > 0$ and

$$P(t) = \left(e^{-\beta\, t}P_0^{-1} + \int_0^t e^{-\beta\,(t-\tau)}\frac{\breve{M}_1^2(\tau)}{N^2(\tau)}\,d\tau\right)^{-1} > 0 \tag{4.86}$$

Therefore, W is a Lyapunov function and $\hat{f}(t)$ globally asymptotically tends to f. □

4.3.2 Adaptive Filters Design

Once the aircraft model (2.42) includes faults on the input sensors, namely on the elevator f_{δ_e}, on the aileron f_{δ_a}, on the rudder f_{δ_r} and on the throttle $f_{\delta_{th}}$ sensors, it is possible to split the overall model into 4 distinct subsystems that can be expressed in the form (4.1). Each of the 4 aircraft models for FDI leads to the form (4.4) by means of a suitable coordinate transformation, as presented in Section 4.1.2. Furthermore, it is straightforward to verify that all the conditions required by Remark 15 are satisfied. Hence, a set of 4 NLGA adaptive filters is designed in the general form (4.73)–(4.75). This scheme allows to estimate the magnitude of a step fault acting on a single input sensor.

Elevator Adaptive Filter

For the FDI aircraft model with the fault on the elevator, the state scalar component $\bar{x}_{1s}$ needed to detect f_{δ_e} is $\bar{x}_{11}$ expressed by (4.12). Hence, it is possible to specify the particular expression of the faulty dynamics (4.72). The design of the NLGA adaptive filter (4.73)–(4.75) for f_{δ_e} is based on these dynamics

$$\begin{aligned}
\dot{\bar{y}}_{1s,e} &= M_{1e}\cdot f_{\delta_e} + M_{2e}\\
M_{1e} &= -\frac{C_{\delta_e}}{mt_d}V^2\\
M_{2e} &= \frac{V^2}{m}\Bigg(-\left(C_{D0} + C_{D\alpha}\alpha + C_{D\alpha_2}\alpha^2\right)\cos\alpha\\
&\quad + \left(C_{L0} + C_{L\alpha}\alpha\right)\sin\alpha - \frac{\left(C_{m0} + C_{m\alpha}\alpha + C_{mq}q_\omega\right)}{t_d}\Bigg)\\
&\quad - g\sin\theta - V\sin\alpha\ q_\omega - \frac{(I_z - I_x)}{mt_d}p_\omega r_\omega - \frac{C_{\delta_e}}{mt_d}V^2\,\delta_e
\end{aligned} \tag{4.87}$$

with $M_{1e}(t) \neq 0$, $\forall\ t \geq 0$.

Aileron Adaptive Filter

For the FDI aircraft model with the fault on the aileron, the state scalar component $\bar{x}_{1s}$ needed to detect f_{δ_a} is $\bar{x}_{11}$ expressed by (4.17). Hence, it is possible to specify the particular expression of the faulty dynamics (4.72). The design of the NLGA adaptive filter (4.73)–(4.75) for f_{δ_a} is based on these dynamics

$$\begin{aligned} \dot{\bar{y}}_{1s,a} &= M_{1a} \cdot f_{\delta_a} + M_{2a} \\ M_{1a} &= \frac{C_{\delta_a}}{I_x} V^2 \\ M_{2a} &= \frac{(C_{l\beta}\beta + C_{lp} p_\omega)}{I_x} V^2 + \frac{(I_y - I_z)}{I_x} q_\omega r_\omega + \frac{C_{\delta_a}}{I_x} V^2 \delta_a \end{aligned} \tag{4.88}$$

with $M_{1a}(t) \neq 0, \forall\, t \geq 0$.

Rudder Adaptive Filter

For the FDI aircraft model with the fault on the rudder, the state scalar component $\bar{x}_{1s}$ needed to detect f_{δ_r} is $\bar{x}_{11}$ expressed by (4.22). Hence, it is possible to specify the particular expression of the faulty dynamics (4.72). The design of the NLGA adaptive filter (4.73)–(4.75) for f_{δ_r} is based on these dynamics

$$\begin{aligned} \dot{\bar{y}}_{1s,r} &= M_{1r} \cdot f_{\delta_r} + M_{2r} \\ M_{1r} &= \frac{C_{\delta_r}}{I_z} V^2 \\ M_{2r} &= \frac{(C_{n\beta}\beta + C_{nr} r_\omega)}{I_z} V^2 + \frac{(I_x - I_y)}{I_z} p_\omega q_\omega + \frac{C_{\delta_r}}{I_z} V^2 \delta_r \end{aligned} \tag{4.89}$$

with $M_{1r}(t) \neq 0, \forall\, t \geq 0$.

Throttle Adaptive Filter

For the FDI aircraft model with the fault on the throttle, the state scalar component $\bar{x}_{1s}$ needed to detect $f_{\delta_{th}}$ is $\bar{x}_{15}$ expressed by (4.27). Hence, it is possible to specify the particular expression of the faulty dynamics (4.72). The design of the NLGA adaptive filter (4.73)–(4.75) for $f_{\delta_{th}}$ is based on these dynamics

$$\begin{aligned} \dot{\bar{y}}_{1s,th} &= M_{1th} \cdot f_{\delta_{th}} + M_{2th} \\ M_{1th} &= \frac{t_f}{n_e} (t_0 + t_1\, n_e) \\ M_{2th} &= t_n\, n_e^3 + \frac{t_f}{n_e} (t_0 + t_1\, n_e)\, \delta_{th} \end{aligned} \tag{4.90}$$

with $M_{1th}(t) \neq 0, \forall\, t \geq 0$.

Remark 16. *The full structure of the NLGA–AF is obtained by replacing the specific expressions of M_{1x}, M_{2x} and $\bar{y}_{1s,x}$, for each subscript $x \in \{e, a, r, th\}$, given by (4.87), (4.88), (4.89) and (4.90) into the general form of the adaptive filter described by the equations (4.73), (4.74) and (4.75).*

4.4 NLGA–PF FDI Scheme

This section addresses the FDI problem for a nonlinear stochastic dynamic system. When stochastic systems are considered, much of the FDI schemes has relied on the system being linear and the noise and disturbances being Gaussian. In such cases, the Kalman filter is usually employed for state estimation and its innovation is then used as the residual (Chen and Patton 1999). The idea used in the linear case mentioned above has been extended to some nonlinear stochastic systems with additive Gaussian noise and disturbance by employing the linearisation and Gaussianization techniques, and in this case, the Kalman filter is usually replaced by the Extended Kalman Filter (EKF) (Doucet *et al.* 2001). Although this EKF–based approach appears straightforward, there are no general results to guarantee that such approximation will work well in most case. The FDI problems in general nonlinear non–Gaussian stochastic systems are still open.

Recently, the Particle Filter (PF), a Monte Carlo based method for nonlinear non–Gaussian state estimation, has attracted much attention (Doucet *et al.* 2001, Zhang *et al.* 2005). Polynomial extended Kalman filters and Unscented Kalman Filters (UKF) represent alternative techniques with performance superior to that of the EKF (Germani *et al.* 2007). However, the interest for PF–based methods stems from their ability of being able to handle any functional nonlinearity and system or measurement noise of any probability distribution. As an example, the work (Zhang *et al.* 2005) represents an attempt to introduce PF into the field of FDI. The fault isolation problem is also investigated.

By combining PF with the NLGA design technique, a particle filtering based approach, *i.e.* the NLGA–PF, to FDI is presented. In particular, the PF is employed to develop a method for solving the FDI problem for the nonlinear stochastic model of the system under diagnosis, which is derived by following a NLGA strategy. The use of the NLGA allows to easily obtain disturbance decoupled residual generators in a stochastic framework. The fault isolation and the disturbance decoupling suggested in this section is different from the method presented in (Zhang *et al.* 2005), as achieved via the NLGA strategy.

4.4.1 Basic Particle Filter Theory

In the following, a short introduction to the basic particle filter theory, also known as bootstrap filter, is provided. For more complete presentations and details, the readers are referred to (Doucet *et al.* 2001).

The general nonlinear discrete–time system in the fault–free case is considered in the

form

$$\begin{aligned} x_{k+1} &= f_d(x_k, c_k) + v_k^x \\ y_k &= g_d(x_k, c_k) + v_k^y \end{aligned} \tag{4.91}$$

where $x_k \in \mathcal{X} \subset \mathcal{R}^{\ell_n}$ is the discrete–time state vector, $c_k \in \mathcal{R}^{\ell_c}$ is the sampled input vector, $y_k \in \mathcal{R}^{\ell_m}$ is the sampled output vector, $v_k^x \in \mathcal{R}^{\ell_n}$ and $v_k^y \in \mathcal{R}^{\ell_m}$ are state and output noises. $f_d(x, c)$ and $g_d(x, c)$ are nonlinear functions. The noise processes v_k^x and v_k^y are assumed to be white with known Probability Density Functions (PDF) $p_x(v_k^x)$ and $p_y(v_k^y)$. The PDF of the initial state x_0 is assumed to be $p_0(x)$. Denote also by $\mathcal{D}_k$ the input–output sampled data observed up to the time instant k, *i.e.* $\mathcal{D}_k = \{(c_i, y_i) : i = 1, \ldots, k\}$.

The filtering problem is to estimate the distribution of the state vector at each instant k, based on the data observed up to instant k, or more precisely, to estimate the conditional PDF $p(x_k|\mathcal{D}_k)$. In general, no accurate finite dimensional filter exists for nonlinear systems, even if the noises are assumed to be Gaussian. The basic idea of PF is to approximate the PDF of the state vector x_k at each instant k with the sum of (a large number of) Dirac functions, and to make them evolve at each time instant based on the latest observed data. Each Dirac function used in the PDF approximation is called a particle.

To start the particle filter at the initial instant $k = 0$, randomly draw M points in $\mathcal{R}^{\ell_n}$ following the assumed PDF $p_0(.)$ of the initial state vector. These M points are denoted with the vectors $\eta_0^j \in \mathcal{R}^{\ell_n}$, $j = 1, \ldots, M$, then $p_0(.)$ is approximated by the relation

$$p\left(x_0|\mathcal{D}_0\right) \approx \frac{1}{M} \sum_{j=1}^{M} \delta\left(x_0 - \eta_0^j\right) \tag{4.92}$$

Recursively, at each instant $k \geq 0$, with

$$p\left(x_k|\mathcal{D}_k\right) \approx \frac{1}{M} \sum_{j=1}^{M} \delta\left(x_k - \eta_k^j\right) \tag{4.93}$$

already estimated, the distribution of x_{k+1} is first predicted with the state equation of the system (4.91), leading to an approximation of the PDF $p(x_{k+1}|\mathcal{D}_k)$. For this purpose, each particle η_k^j, for $j = 1, \ldots, M$, is propagated following the state equation of the system (4.91) to the position $f_d(\eta_k^j, c_k)$ and perturbed by a random vector γ_k^j drawn following the state noise PDF $p_x(.)$, and allowing the computation of

$$\eta_{k+1|k}^j = f_d(\eta_k^j, c_k) + \gamma_k^j \tag{4.94}$$

Then

$$p(x_{k+1}|\mathcal{D}_k) \approx \frac{1}{M} \sum_{j=1}^{M} \delta(x_{k+1} - \eta_{k+1|k}^j) \tag{4.95}$$

Now the data observed at instant $k+1$ are used to estimate $p(x_{k+1}|\mathcal{D}_{k+1})$. According to the Bayes rule, each particle $\eta^j_{k+1|k}$ is weighted by its likelihood w^j_{k+1} based on the output equation of the system (4.91), the following relations hold

$$\begin{aligned} w^j_{k+1} &= p_y\left(y_{k+1} - g_d(\eta^j_{k+1|k}, c_k)\right) \\ S_{k+1} &= \sum_{j=1}^{M} w^j_{k+1} \\ p(x_{k+1}|\mathcal{D}_{k+1}) &\approx \frac{1}{S_{k+1}} \sum_{j=1}^{M} w^j_{k+1}\delta(x_{k+1} - \eta^j_{k+1|k}) \end{aligned} \tag{4.96}$$

In order to approximate $p(x_{k+1}|D_{k+1})$ with M equally weighted particles, M points are randomly drawn following the discrete probability distribution in the form

$$P(x = \eta^j_{k+1|k}) = \frac{w^j_{k+1}}{S_{k+1}}\,, \qquad j = 1, \ldots, M \tag{4.97}$$

The resulting points, noted as $\eta^j_{k+1} \in \mathcal{R}^{\ell_n}$ for $j = 1, \ldots, M$, are then used to make the following approximation

$$p(x_{k+1}|\mathcal{D}_{k+1}) \approx \frac{1}{M} \sum_{j=1}^{M} \delta(x_{k+1} - \eta^j_{k+1}) \tag{4.98}$$

The algorithm then goes to the next iteration with k increased by 1.

The software code for the implementation of the PF strategy (Doucet *et al.* 2001, Zhang *et al.* 2005) is freely availabe at the website `http://www.cs.ubc.ca/~nando/software.html`.

4.4.2 Throttle Particle Filter Design

In the following the particle filter to detect the throttle sensor fault is designed. The design of the particle filters related to the remaining input sensors is not presented in this work.

As for the NLGA and the NLGA-AF, the NLGA–PF has to be designed from the $\bar{x}_1$ subsystem (4.4). However, as the PF algorithm requires a discrete–time system, the following model in the form (4.91) with $\ell_n = \ell_c = \ell_m = 1$ is derived by using the simple Euler forward discretisation method, with a sampling time of 0.01 s

$$\begin{aligned} \xi_{k+1} &= \xi_k + 0.01\left(t_n\xi_k^3 + \frac{t_f}{\xi_k}(t_0 + t_1\xi_k)\delta_{thk}\right) + \zeta_k \\ y_k &= \xi_k + \nu_k \end{aligned} \tag{4.99}$$

The scalar processes ν_k and ζ_k describe the measurement noise and the effect of the non critical disturbances, respectively, whilst δ_{thk} and y_k are the sampled input–output data sequences. Finally, the FDI residuals of the NLGA–PF are computed as the difference between the sampled data n_e and its prediction provided by the PF.

Remark 17. *As shown in Section 4.2, the NLGA filters with robustness improvement are structurally decoupled from critical disturbance and optimised in order to maximise the fault sensitivity with respect to non critical disturbances. Thus, the NLGA filters are suitable to be exploited in a stochastic framework and can be compared with the NLGA–PF.*

Chapter 5

Simulation Results

The simulation results obtained by means of the Matlab/Simulink® aircraft simulator are summarised in this chapter. The adopted residual evaluation logic is explained in Section 5.1. Section 5.2 describes the FDI problem for a complete aircraft trajectory formed by a prescribed set of steady–state flight condition. Sections 5.3 show basic simulation results and performances evaluation. In Section 5.4 the proposed PM and NLGA techniques are compared with other FDI schemes and the robustness with respect to a complete aircraft trajectory is evaluated. Finally, in order to evaluate robustness with respect to uncertainty acting on the system, a Monte–Carlo analysis is performed in Section 5.5.

5.1 Threshold Test

Once the residuals have been generated, the residual evaluation logic is used to detect and isolate any fault occurrence. The residual processing methods can be based on simple residual geometrical analysis or comparison with fixed thresholds (Chen and Patton 1999). More complex residual evaluation can rely on statistical properties of the residual and hypothesis testing (Basseville and Nikiforov 1993), or based on adaptive threshold, that is, the so–called threshold selector (Emami-Naeini *et al.* 1988).

In general, in the absence of faults, the residual signals are approximately zero. In practical situations, the residual is never zero, even if no faults occur. A threshold must then be used and normally is set suitably larger than the largest magnitude of the residual for the fault-free case. The smallest detectable fault is a fault which drives the residual function to just exceed the threshold. Any fault producing a residual response smaller than this magnitude is not detectable.

More in detail, the most widely used way to fault detection is achieved by directly comparing residual signal $r(t)$ or a residual function $J(r(t))$ with a fixed threshold ε or a threshold function $\varepsilon(t)$ as follows

$$\begin{aligned} J(r(t)) &\leq \varepsilon(t) \quad \text{for } f(t) = 0 \\ J(r(t)) &> \varepsilon(t) \quad \text{for } f(t) \neq 0 \end{aligned} \tag{5.1}$$

where $f(t)$ is the general fault vector. If the residual exceeds the threshold, a fault may be occurred. This test works especially well with fixed thresholds ε if the process operates approximately in steady–state and it reacts after relatively large feature, *i.e.* after either a large sudden or a long-lasting gradually increasing fault.

In practice, if the residual signal is represented by the stochastic variable $r(t)$, mean value and variance are computed as follows

$$\begin{aligned} \bar{r} &= E\{r(t)\} = \frac{1}{N}\sum_{t=1}^{N} r(t) \\ \sigma_r^2 &= E\{(r(t)-\bar{r})^2\} = \frac{1}{N}\sum_{t=1}^{N}(r(t)-\bar{r}^2) \end{aligned} \tag{5.2}$$

where $\bar{r}$ and σ_r^2 are the normal values for the mean and variance of the fault-free residual, respectively. N is the number of samples of the vector $r(t)$. Therefore, the threshold test for FDI of (5.1) is rewritten as

$$\begin{aligned} &\bar{r} - \nu\,\sigma_r \le r(t) \le \bar{r} + \nu\,\sigma_r && \text{for } f(t) = 0 \\ &r(t) < \bar{r} - \nu\,\sigma_r \quad \text{or} \quad r(t) > \bar{r} + \nu\,\sigma_r && \text{for } f(t) \neq 0 \end{aligned} \tag{5.3}$$

i.e. the comparison of $r(t)$ with respect to its statistical normal values. In order to separate normal from faulty behaviour, the tolerance parameter ν (normally $\nu \ge 3$) is selected and properly tuned. Hence, by a proper choice of the parameter ν, a good trade–off can be achieved between the maximisation of fault detection probability and the minimisation of false alarm probability.

In practice, the threshold values depend on the residual error amount due to the measurements errors, the model approximations and the disturbance signals that are not completely decoupled.

5.2 FDI Procedure for a Complete Trajectory

The target of the proposed FDI schemes is to perform the aircraft fault diagnosis in a prescribed set of steady–state flight conditions, which cover the largest part of the complete trajectory. Each of these steady–state flight conditions can be described by both its trim point and a mathematical model. Hence, it is possible to perform the off–line design of a set of residual generators for each of these flight conditions.

In the considered framework, a simple FMS (Flight Management System) (Collinson 2002) is supposed installed on board, and its main tasks are:

- Scheduling the current reference flight condition, since the whole trajectory, defining the flight plan, is described by a sequence of steady–state flight conditions.
- Computing an accurate navigation solution exploiting the sensor measurements.

- Providing to the FDI subsystem the time intervals corresponding to an aircraft state sufficiently near to the current reference flight condition, so that it is possible to apply the proper residual generator filters.

Remark 18. *The set of all the allowed steady–state flight conditions can be parameterised (speed, radius of curvature and flight–path angle) on a manifold and there exist bijective functions mapping both to the input trim manifold and to the output trim manifold. As a consequence the FMS is able to determine when the aircraft motion can be considered sufficiently near to the steady–state condition either by monitoring the input and the output data independently, even if a single fault occurs.*

On the basis of the previous considerations, a possible implementation of the FDI procedure for a complete trajectory could consist of the following steps:

1. Off–line design and optimisation of the residual generators for each trajectory elementary path (high computational cost, but performed off–line).

2. On–line steady–state flight condition recognition by the FMS (task requiring a low computational cost).

3. Switching to the corresponding stored residual generators on the basis of the current working condition.

5.3 Simulations and Performances Evaluation

The chosen single steady–state flight condition for the design of both the PM and the NLGA–based residual generators is a coordinated turn at constant altitude characterised as follows:

- The true air speed is 50 m/s.
- The curvature radius is 1000 m.
- The flight–path angle is 0^o.
- The altitude is 330 m.
- The flap deflection is 0^o.

This represents one of most general flight condition due to the coupling of the longitudinal and lateral dynamics. Moreover, it is used in simulation to highlight the performances of the proposed methods in the nominal flight condition.

5.3.1 PM

The PM residual generator filters are fed by the 4 component input vector $c(t)$ and the 9 component output vector $y(t)$ acquired from the nonlinear simulation aircraft model described in Chapter 2. In particular, as presented in Section 3.3, a bank of 4 residual generator filters has been used to detect input sensor faults regarding the 4 input variables $c(t) = \left[\Delta\delta_e(t) \quad \Delta\delta_a(t) \quad \Delta\delta_r(t) \quad \Delta\delta_{th}(t)\right]^{\mathrm{T}}$. Moreover, in order to obtain the fault isolation properties, each residual generator function of the input bank is fed by all but one the 4 input signals and by the 9 output variables $y(t) = \left[\Delta V(t) \quad \Delta p_\omega(t) \quad \Delta q_\omega(t) \quad \Delta r_\omega(t) \quad \Delta\phi(t) \quad \Delta\theta(t) \quad \Delta\psi(t) \quad \Delta H(t) \quad \Delta n_e(t)\right]^{\mathrm{T}}$.

Remark 19. *The measurements of $\alpha(t)$ and $\beta(t)$ have not been considered for FDI, because the structural detectability conditions are anyway fulfilled. Moreover, as described in Section 2.2.4, the sensor package provides the value of the variables in $y(t)$ by processing several measurements. However, this situation is not critical for the residual generators (3.2). In fact, due to the assumptions regarding the IMU and the HRS, a fault regarding a single sensor affects only a component of the output vector $y(t)$. Moreover, thanks to the different features of the gyroscope units, system stability and performance are not affected.*

Each filter of the input bank is independent of one of the 4 input signals and then is also insensitive to the corresponding fault signals. Obviously, the residual generator banks have been designed to be decoupled from 3 wind gust signals $d(t) = \left[w_u(t)\, w_v(t)\, w_w(t)\right]^{\mathrm{T}}$, which represent disturbance terms acting on the aircraft system. The capabilities of the FDI system are hence related to the properties of the residual generator functions in the presence of measurement errors, modelling approximations and disturbance signals that cannot be completely decoupled.

The robustness properties of the filters in terms of fault sensitivity and disturbance insensitivity can be achieved according to Section 3.2. The synthesis of the dynamic filters for FDI has been performed by choosing a suitable linear combination of residual generator functions. This choice has to maximise the steady–state gain of the transfer functions shown by (3.48) between input sensor fault signals $f_{c_i}(t)$ and residual functions $r_{c_j}(t)$. Moreover, for each residual generator, the roots of the polynomial matrix $R_{c_j}(s)$ have been optimised and placed in a range between -1 and -10^{-2} for maximising the fault detection promptness, as well as to minimise the occurrence of false alarms.

In the same way, an appropriate filter bank for the output sensor fault isolation, generating the 9 residual functions $r_{o_j}(t)$, has been designed.

In order to assess the diagnosis technique, different fault sizes have been simulated on each sensor. Single faults in the input–output sensors have been generated by producing positive and negative abrupt (step) variations in the input–output signals $c(t)$ and $y(t)$.

The residual signals indicate fault occurrence according to whether their values are lower or higher than the thresholds fixed in fault–free conditions. As described by (5.3), the threshold values depend on the residual error amount due to measurement errors,

linearised model approximations and disturbance signals that are not completely decoupled. A suitable value of $\nu = 4$ for the computation of the positive and negative threshold in (5.3) has been considered, in order to minimise the false alarm occurrence and to maximise the fault sensitivity.

As an example, the 4 residual functions $r_{c_j}(t)$ generated by the filter bank for input sensor fault isolation, under both fault–free and faulty condition are shown in Figure 5.1. Continuous lines represent the fault–free residual functions, while the dotted lines depict the faulty residual signals. Moreover horizontal lines represents the thresholds. The fault has been generated on the 1–st input sensor of the considered aircraft, starting at time $t = 150\,\mathrm{s}$.

The 1–st residual function of Figure 5.1 provides also the isolation of a fault regarding the considered input sensor $f_{c_1}(t)$. It does not depend on a fault affecting the input sensor itself, as the corresponding residual $r_{c_1}(t)$ filter has been designed to be sensitive to the input signal $c^{*1}(t)$.

In a similar way, Figure 5.2 shows the 9 residual functions $r_{o_j}(t)$ generated by the filter bank for output sensor fault isolation, under both fault–free and faulty conditions.

Figures 5.1 and 5.2 show also the ranges out of which the input and output sensor faults are detectable. These maximal and minimal values assumed by the $r_{c_j}(t)$ and $r_{o_j}(t)$ functions in fault–free conditions must be computed with acceptable false–alarms rates.

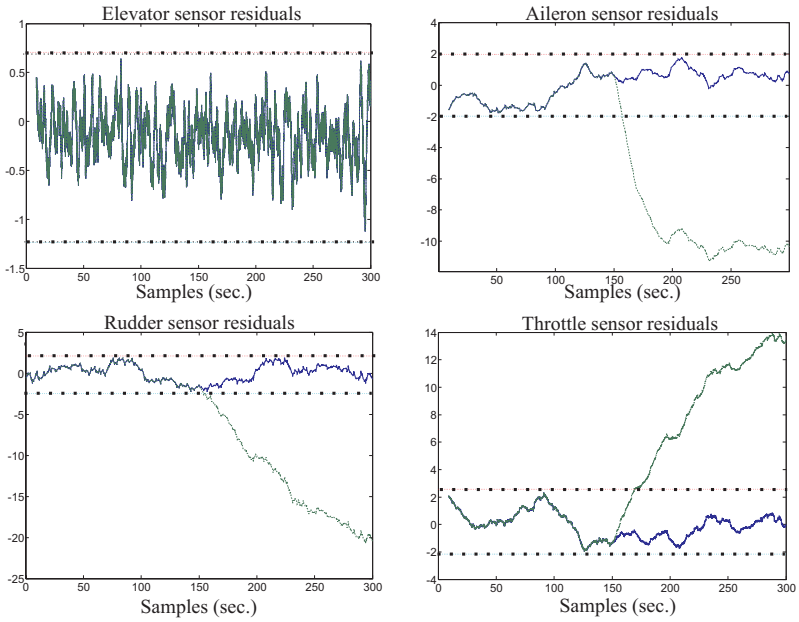


Figure 5.1: Bank residuals for the 1–st input sensor fault isolation.

To summarise the performance of the FDI technique, the minimal detectable step fault amplitudes on the various input and output sensors with the related detection delay times are collected in Table 5.1 and Table 5.2, respectively.

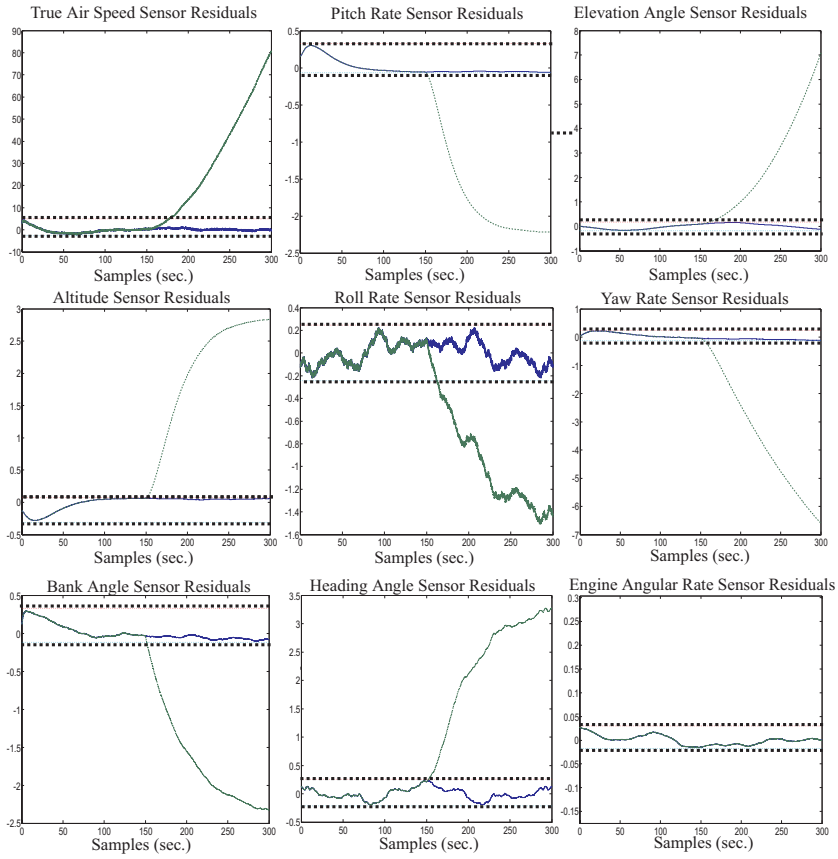


Figure 5.2: Bank residuals for the 9–th output sensor fault isolation.

The minimal detectable fault values in Tables 5.1 and 5.2 are expressed in the unit of measure of the sensor signals. The fault sizes are relative to the case in which the occurrence of a fault is detected and isolated as soon as possible.

The detection delay times, reported in Tables 5.1 and 5.2 represent the worst case results. They are evaluated on the basis of the time taken by the slowest residual function, or by the estimation of a fault, to cross the settled threshold.

Table 5.1: PM minimal detectable step input sensor faults.

Sensor $c_i(t)$	**Var.**	**Fault Size**	**Delay**
Elevator deflection angle	δ_e	2^o	18 s
Aileron deflection angle	δ_a	3^o	6 s
Rudder deflection angle	δ_r	4^o	8 s
Throttle aperture %	δ_{th}	2%	15 s

Table 5.2: PM minimal detectable step output sensor faults.

Sensor $y_i(t)$	**Var.**	**Fault Size**	**Delay**
True Air Speed	V	8 m/s	9 s
Pitch Rate	q_ω	3°/s	22 s
Elevation Angle	θ	5°	10 s
Altitude	H	8 m	12 s
Roll Rate	p_ω	2°/s	24 s
Yaw Rate	r_ω	3°/s	29 s
Bank Angle	ϕ	5°	5 s
Heading Angle	ψ	6°	20 s
Engine Speed	n_e	20 rpm	25 s

Remark 20. *With reference to the application domain of general aviation aircrafts, the severity of each fault condition can be classified. The considered fault conditions can be ordered as follows, from the most to the least critical variable:*

- δ_e, δ_r, δ_a *and* δ_{th}.
- V, ϕ, θ *and* n_e.
- ψ *and* H.
- p_ω, q_ω *and* r_ω.

The main criterion used to state the severity list is based on the dynamics of the monitored variables. In particular, the faster the time scale of a variable, the greater the severity of the associated fault. However, faults on the variables p_ω, q_ω *and* r_ω *are the less critical, even if their time scales are not the slowest. In fact, classical autopilots for general aviation aircrafts usually do not exploit these measurements very much. Moreover, feedback control schemes adopting high–gain with respect to the angular rate components are typically used only if the modes of the aircraft dynamics need to be drastically changed in order to fulfil the required flying qualities.*

On the basis of the severity list, the FDI filter optimisation described here has been hence performed in order to enhance the FDI of the most critical measurement sensors, i.e. *for optimising the related fault sensitivity and detection delay time.*

5.3.2 NLGA

The NLGA–based FDI schemes presented in Chapter 4 have been designed as follows:

NLGA design. A bank of 4 filters has been used in order to perform the fault diagnosis and isolation on the input sensors. The filters are designed as described in Section 4.1. The synthesis of the filters has been performed by using filter gains

that optimise the fault sensitivity and reduce as much as possible the occurrence of false alarms due to model uncertainties and to disturbances not completely decoupled. This robustness requirement has been fulfilled by designing the residual gains according to the procedure described in Section 4.2. For example, with reference to the fourth residual generator, this procedeure has led to $k_{\delta_{th}} = 1$ which satisfies the norm bounds $\gamma = 1.2$ and $\beta = 400$. This guarantees a good separation on the residual signal with $\|f\|_{\mathcal{L}_2} \geq 0.05$ and $\|d\|_{\mathcal{L}_2} \leq 10$, where $\mathcal{L}_2$-norm is considered.

NLGA–AF design. A bank of 4 adaptive filters has been used in order to perform the diagnosis and isolation as well as the estimation of the f_{δ_e}, f_{δ_a}, f_{δ_r} and $f_{\delta_{th}}$ fault size. The adaptive filter designs have been carried out according to the procedure described in Section 4.3.

NLGA–PF design. The filter for the FDI of throttle sensor is implemented via the algorithm summarised in Section 4.4 with a number $M = 200$ particles and it uses 20000 sampled data δ_{thk} and n_{ek}, acquired from the continuous–time aircraft model (2.42). Moreover, the PDF for the stochastic processes affecting the system (4.91) are easily estimated from the mathematical knowledge of the aircraft flight simulator and its measurements (see Section 2.2.4). It is worth noting that in this case the isolation of the throttle actuator fault is enhanced, since the scalar $\bar{x}_1$–subsystem (4.4) is affected by a single sensor fault and it is decoupled from the faults affecting the remaining sensors (elevator, aileron, and rudder). The scalar structure of the $\bar{x}_1$–subsystem (4.4) facilitates also the optimal choice of the parameters for the construction of the PF (Zhang *et al.* 2005), and improves the Sampling Importance Resampling (SIR) strategy selected for posterior PDF estimation and the importance weights defined in Section 4.4 (Doucet *et al.* 2001).

Each filter obtained by the described design procedures is structurally decoupled from the vertical and lateral wind disturbance components and is sensitive to a single input sensor fault.

Remark 21. *For the proposed application, the NLGA–based FDI schemes considers only the faults on the inputs sensors. In fact, the output sensor faults cannot be directly modelled by (4.1). On the other hand, a fault modelled by means of an augmented state, as reported in (Massoumnia 1986, Zad and Massoumnia 1999), leads to a nonlinear system which does not fulfill the structural fault detectability condition $\ell(x) \notin \Omega^*$.*

In order to asses the NLGA diagnosis techniques, in similar way to PM evaluation, single steps faults have been used. Moreover, also in this case, the threshold values have been experimentally chosen according to (5.3). A suitable value of $\nu = 8$ for the computation of the positive and negative threshold in (5.3) has been considered.

In Figures 5.3 and 5.4, the simulation results referring to a particular case are reported, where a small step fault f_{δ_e} with size of 2^o starting at time $t = 150\,\mathrm{s}$ is added to the elevator sensor. Figures 5.3 and 5.4 are referred to the NLGA and NLGA–AF, respectively. The

behaviour of r_{δ_e} and $\hat{f}_{\delta_e}$ highlights a better detection time than the corresponding one of the PM. Moreover, the remaining residuals r_{δ_a}, r_{δ_r}, $r_{\delta_{th}}$ and estimates $\hat{f}_{\delta_a}$, $\hat{f}_{\delta_r}$, $\hat{f}_{\delta_{th}}$ never cross the corresponding thresholds, so that the fault isolation is achieved. Note that the estimate $\hat{f}_{\delta_e}$ is accurate, even with a small fault size.

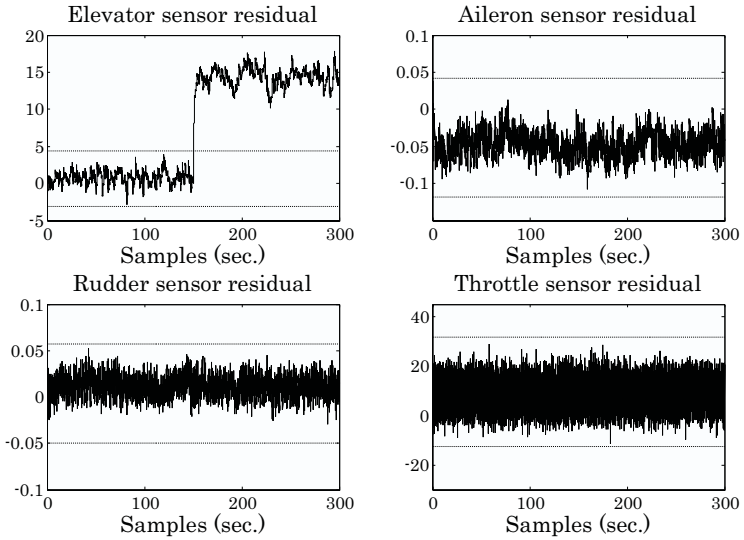


Figure 5.3: NLGA elevator sensor FDI.

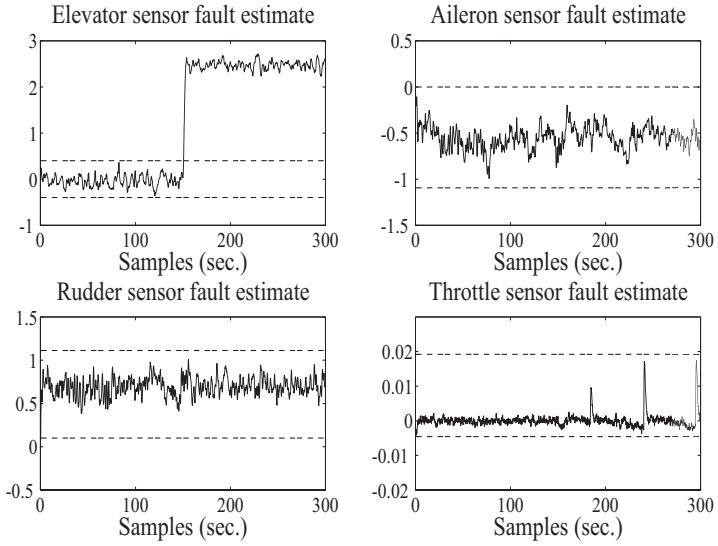


Figure 5.4: NLGA–AF elevator sensor FDI with fault size estimation.

As an example, the residual generated by the NLGA and the NLGA–PF for the δ_{th} sensor FDI, under both fault–free and faulty condition, are shown in Figure 5.5. Continuous line represent the fault free residual functions, while the dotted lines depicts the faulty residual signals. The fault has been generated on the throttle sensor of the considered aircraft, starting at time $t = 100\,\mathrm{s}$.

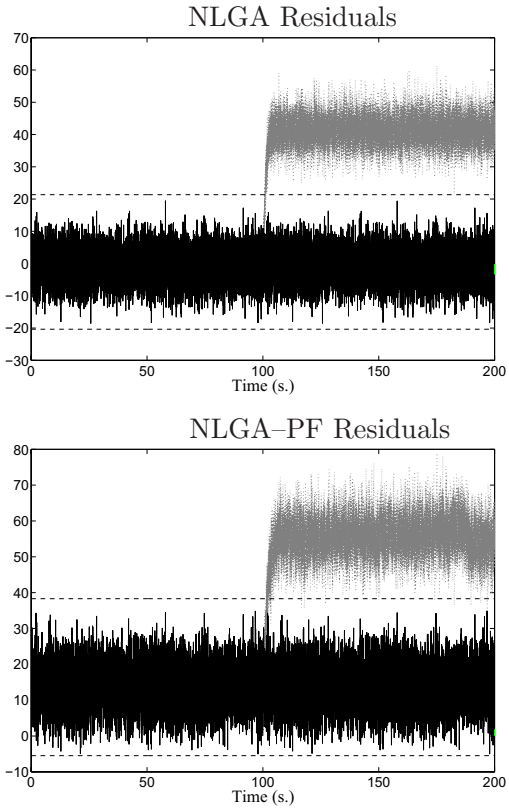


Figure 5.5: NLGA and NLGA–PF residuals for throttle sensor FDI.

In order to summarise the performance of the proposed NLGA, NLGA–AF and NLGA–PF FDI schemes, the minimal detectable step fault amplitudes on the various input sensors with the related detection delay time are collected in Tables 5.3, 5.4 and 5.5, respectively.

Table 5.3: NLGA minimal detectable step input sensor faults.

Sensor $c_i(t)$	**Var.**	**Fault Size**	**Delay**
Elevator deflection angle	δ_e	2^o	5 s
Aileron deflection angle	δ_a	2^o	3 s
Rudder deflection angle	δ_r	2^o	6 s
Throttle aperture %	δ_{th}	6%	3 s

Table 5.4: NLGA-AF minimal detectable step input sensor faults.

Sensor $c_i(t)$	**Var.**	**Fault Size**	**Delay**
Elevator deflection angle	δ_e	2^o	6 s
Aileron deflection angle	δ_a	2.5^o	4 s
Rudder deflection angle	δ_r	4^o	6 s
Throttle aperture %	δ_{th}	5%	5 s

Table 5.5: NLGA-PF minimal detectable step throttle sensor fault.

Sensor $c_i(t)$	**Var.**	**Fault Size**	**Delay**
Throttle aperture %	δ_{th}	3%	3 s

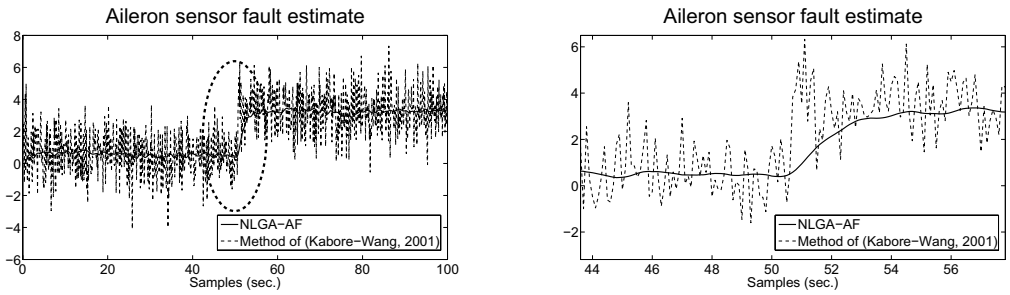
Remark 22. *For the considered aircraft application, the computational burden of NLGA and NLGA–AF is lower than that of NLGA–PF, so that they are suitable for low–cost implementations. On the other hand, the NLGA–PF provides the minimal detectable fault size.*

Remark 23. *A peculiarity of NLGA–AF is that it provides not only FDI but also a fault estimate. For this reason, it is useful to evaluate it in comparison with the fault identification scheme proposed in (Kaboré and Wang 2001, Kaboré* et al. *2000). In particular, in the considered aircraft application, a fault estimator for the aileron input sensor can be easily derived according to the procedure described in (Kaboré and Wang 2001), exploiting the expression of the roll rate p_ω dynamic equation. In Figure 5.6, it is shown the result of a simulation in which a fault of* 2.5^o *is affecting the aileron input sensor. As can be seen, the proposed NLGA–AF is less sensitive to measurement noise, which allows to obtain a smaller minimal detectable fault. On the other hand, the fault estimation technique of (Kaboré and Wang 2001) provides a faster response and, therefore, a slower detection time.*

Advantages and drawbacks of the PM and the NLGA–based FDI methods developed in this work can be summarised as follows:

- Both PM filters and NLGA perform low-pass filtering of input/output measurements. PM by means of the poles of $R(s)$, designed according to an off-line optimisation procedure; NLGA by means of first-order low pass filters. However, the degree of $R(s)$ is generally greater than 1, so the filtering action of PM can be more efficient.

- For the considered aircraft application, the computational burden of polynomial filters is lower than that of NLGA filters, so that they are suitable for low-cost implementations.



(a) Full plot (b) Detailed plot

Figure 5.6: Comparison between NLGA–AF and fault estimator of Kaboré and Wang.

- NLGA can obtain smaller detection time, compared with PM filters, thanks to the fact that they directly take into account nonlinear terms.

5.4 Comparative Studies and Robustness Evaluation

In this section, the robustness characteristics of the proposed PM and NLGA FDI schemes have been evaluated and compared also with respect to the UIKF (Unknown Input Kalman Filter) scheme (Chen and Patton 1999) and the NN (Neural Nets) technique (Korbicz *et al.* 2004). The robustness is achieved by using the same residual generators for a large set of flight condition. In the following a brief description of the adopted design procedure for the UIKF and NN FDI schemes is provided:

UIKF desgin. A bank of UIKF has been exploited for diagnosing faults of the monitored process. This technique seems to be robust with respect to the modelling uncertainties, the system parameter variations and the measurement noise, which can obscure the performance of a FDI system by acting as a source of false faults. The procedure recalled here requires the design of an UIKF bank and the basic scheme is the standard one: a set of measured variables of the system is compared with the corresponding signals estimated by filters to generate residual functions. The diagnosis has been performed by detecting the changes of UIKF residuals caused by a fault. The FDI input sensor scheme exploits a number of KF equal to the number of input variables. Each filter is designed to be insensitive to a different input sensor of the process and its disturbances (the so–called unknown inputs). Moreover, the considered UIKF bank was obtained by following the design technique described in (Chen and Patton 1999) (Section 3.5, pp. 99–105), whilst the noise covariance matrices were estimated as described in (Simani *et al.* 2002) (Section 3.3, pp. 70–74 and Section 4.6, pp. 130–131). Each of the 4 UIKF of the bank was de–coupled from both one input sensor fault and the wind gust disturbance component, thus providing the optimal filtering of the input–output measurement noise sequences.

NN design. A dynamic NN bank has been exploited in order to find the dynamic connection from a particular fault regarding the input sensors to a particular residual. In this case, the learning capability of NN is used for identifying the nonlinear dynamics of the monitored plant. The dynamic NN provides the prediction of the process output with an arbitrary degree of accuracy, depending on the NN structure, its parameters and a sufficient number of neurons. Once the NN has been properly trained, the residuals have been computed as the difference between predicted and measured process outputs. The FDI is therefore achieved by monitoring residual changes. The NN learning is typically an off–line procedure. Normal operation data are acquired from the monitored plant and are exploited for the NN training. Regarding the NN FDI method, and according to a Generalised Observer Scheme (GOS) (Chen and Patton 1999), a bank of 4 time–delayed three–layers Multi–Layer Perceptron (MLP) NN with 15 neurons in the input layer, 25 neurons in the hidden layer and 1 neuron in the output layer is implemented. Each NN was designed to be insensitive to each input sensor fault, and the NN were trained in order to provide the optimal output prediction on the basis of the training pattern and target sequences (Korbicz *et al.* 2004).

The performances of the different FDI schemes have been evaluated by considering a more complex aircraft trajectory. This has been obtained by means of the guidance and control functions of a standard autopilot, which stabilises the aircraft motion towards the reference trajectory as depicted in Figure 5.7.

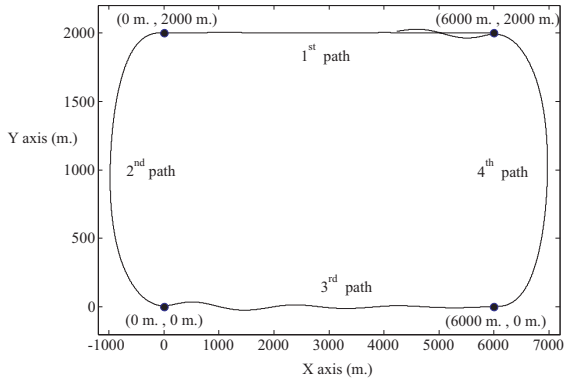


Figure 5.7: Aircraft complete trajectory example.

The reference trajectory is made up of 4 branches (2 straight flights and 2 turn flights) so that a closed path is obtained. It is worth observing that only 2 steady–state flight conditions are used to follow alternatively the 4 branches of the reference trajectory:

- Straight flight condition (1–st and 3–rd path):
  - true air speed = 50 m/s.

  - radius of curvature = ∞.
  - flight–path angle = 0^o.
  - altitude = 330 m.
  - flap deflection = 0^o.

- Turn flight condition (2–nd and 4–th path):
  - true air speed = 50 m/s.
  - radius of curvature = 1000 m.
  - flight–path angle = 0^o.
  - altitude = 330 m.
  - flap deflection = 0^o.

Note that the reference turn flight condition is used to design the PM and the NLGA filters in Section 5.3. The achieved results are reported in Tables 5.1 and 5.3, respectively. The performed tests represent also a possible reliability evaluation of the considered FDI techniques. In fact, in this case the diagnosis requires that the residual generators are robust with respect to the flight conditions that do not match the nominal trajectory used for the design.

Table 5.6 summarises the results obtained by considering the observers and filters (corresponding to the PM, NLGA, UIKF and NN) for the input sensor FDI, whose parameters have been designed and optimised for the steady–state coordinated turn represented by the 2–nd reference flight condition of the complete trajectory. Table 5.6 reports the performances of the considered FDI techniques in terms of the minimal detectable step faults on the various input sensors, as well as the corresponding parameters ν for the residual evaluation of (5.3). The mean detection delay is also reported in Table 5.6 in order to compare the effectiveness of the different FDI schemes.

Table 5.6: Performances for a complete aircraft trajectory.

Variable	**PM**	**NLGA**	**UIKF**	**NN**
ν	4	12	9	5
δ_e	4^o	3^o	4^o	3^o
δ_a	5^o	3^o	5^o	4^o
δ_r	5^o	3^o	4^o	4^o
δ_{th}	7 %	10 %	11 %	12 %
Mean Detection Delay	26 s	25 s	31 s	27 s

The choice of ν has been performed with reference to the particular flight conditions involved in the complete trajectory following. In particular, the selected value of ν for each FDI observer or filter represents a trade–off between two objectives, *i.e.* for increasing

the residual fault sensitivity and promptness, as well as for minimising the occurrence of false alarms due to the switching among the reference flight conditions needed to stabilise the aircraft motion towards the reference trajectory. Table 5.6 shows how the proper design of the parameter ν allows to obtain good performances with all the considered FDI schemes, hence the robustness with respect to the proposed complete trajectory is always achieved.

It is worth noting that the NLGA has a theoretical advantage by taking into account the nonlinear dynamics of the aircraft. However the behaviour of the related nonlinear residual generators is quite sensitive to the model uncertainties due to variation of the flight condition. In fact, the NLGA requires high values of ν which have to be increased (from 8 to 12 in this work) when the aircraft motion regarding the complete trajectory is considered in place of the nominal flight condition. In particular, even though the analysis was restricted just to the aircraft turn phase of the complete trajectory, a performance worsening would happen, since the steady–state condition (nominal flight condition) is quite far to be reached. However, the filter design based on the NLGA lead to a satisfactory fault detection, above all in terms of promptness. On the other hand, regarding the PM, it is rather simple to note the good FDI performances, even if optimisation stages can be required. The ν values selected for the PM are lower, but the related residual fault sensitivities are even smaller. Similar comments can be made for the UIKF and NN techniques.

The simulation model applied to the complete trajectory is an effective way to test the performances of the proposed FDI methods with respect to modelling mismatch and measurement errors. The obtained results demonstrate the reliability of the PM, NLGA, UIKF and NN based FDI schemes as long as proper design procedures are adopted.

5.5 Monte–Carlo Analysis

In this section, further experiment results have been reported. They regard the performance evaluation of the developed FDI schemes with respect to uncertainty acting on the system. Hence, the simulation of different fault–free and faulty data sequences was performed by exploiting the aircraft Matlab/Simulink® simulator and a Monte–Carlo analysis implemented in the Matlab® environment.

The Monte–Carlo tool is useful at this stage as the FDI performances depend on the residual error magnitude due to the system uncertainty, as well as the signal $c(t)$ and $y(t)$ measurement errors. It is worth noting how the Monte–Carlo simulations have been achieved by perturbing the parameters of the PM filter residuals by additive white Gaussian noises with standard deviation values equal to a fixed percentage p of the element values. The same experiments have been performed by statistically varying the main parameters of the NLGA filters. In these conditions, the Monte Carlo analysis represents a further method for estimating the reliability and the robustness of the developed FDI schemes, when applied to the considered aircraft.

For robustness and reliability experimental analysis of the FDI schemes, some perfor-

mance indices have been used. The performances of the FDI method are then evaluated on a number of Monte–Carlo runs equal to 1000. This number of simulations is carried out to determine the indices listed below with a given degree of accuracy:

False Alarm Probability (r_{fa}): the number of wrongly detected faults divided by total fault cases.

Missed Fault Probability (r_{mf}): for each fault, the total number of undetected faults, divided by the total number of times that the fault case occurs.

True Detection/Isolation Probability (r_{td}, r_{ti}): for a particular fault case, the number of times it is correctly detected/isolated, divided by total number of times that the fault case occurs.

Mean Detection/Isolation Delay (τ_{md}, τ_{mi}): for a particular fault case, the average detection/isolation delay time.

These indices are hence computed for the number of Monte–Carlo simulations and for each fault case. Tables 5.7 and 5.8 summarises the results obtained by considering the PM and NLGA, dynamic filters for the input sensor FDI for a complete aircraft trajectory and with $p = 10\%$. The same analysis can be applied again to the residual generated by means of the NN and UIKF FDI schemes, the results are summarised in Tables 5.9 and 5.10.

Table 5.7: PM Monte–Carlo analysis with $\nu = 4$ and $p = 10\%$.

Faulty sensor	r_{fa}	r_{mf}	r_{td}, r_{ti}	τ_{md}, τ_{mi}
δ_e	0.002	0.003	0.997	27 s
δ_a	0.001	0.001	0.999	18 s
δ_r	0.002	0.003	0.997	25 s
δ_{th}	0.003	0.002	0.998	35 s

Table 5.8: NLGA Monte–Carlo analysis with $\nu = 12$ and $p = 10\%$.

Faulty sensor	r_{fa}	r_{mf}	r_{td}, r_{ti}	τ_{md}, τ_{mi}
δ_e	0.003	0.004	0.996	30 s
δ_a	0.002	0.002	0.998	15 s
δ_r	0.001	0.001	0.999	23 s
δ_{th}	0.004	0.003	0.997	32 s

Tables 5.7–5.10 show how the proper design of the dynamic filters with a proper choice of the FDI thresholds allow to achieve false alarm and missed fault probabilities less than

Table 5.9: NN Monte–Carlo analysis with $\nu = 5$.

Faulty sensor	r_{fa}	r_{mf}	r_{td}, r_{ti}	τ_{md}, τ_{mi}
δ_e	0.004	0.005	0.995	33 s
δ_a	0.003	0.003	0.997	23 s
δ_r	0.004	0.004	0.996	29 s
δ_{th}	0.005	0.003	0.997	38 s

Table 5.10: UIKF Monte–Carlo analysis with $\nu = 9$.

Faulty sensor	r_{fa}	r_{mf}	r_{td}, r_{ti}	τ_{md}, τ_{mi}
δ_e	0.003	0.004	0.996	26 s
δ_a	0.002	0.002	0.998	17 s
δ_r	0.001	0.002	0.998	26 s
δ_{th}	0.004	0.003	0.997	37 s

0.6%, detection and isolation probabilities bigger than 99.4%, with minimal detection and isolation delay times. The results demonstrate also that Monte–Carlo simulation is an effective tool for testing and comparing the design robustness of the proposed FDI methods with respect to modelling uncertainty ($p = 10\%$) and fixed measurement errors. This last simulation technique example hence facilitates an assessment of the reliability of the developed, analysed and applied FDI methods.

Chapter 6

Conclusion

The thesis provided theoretical and application results in the detection and isolation of faults on the sensors of a nonlinear aircraft system by using two FDI schemes: the PM and the NLGA. Moreover, two further FDI techniques belonging to the NLGA framework have been developed: the NLGA–AF and the NLGA–PF.

In the following, the main topics and contributions presented in this thesis are summarised chapter by chapter:

- Chapter 1 has presented an introduction to the fault diagnosis problem and the most popular FDI approaches were briefly recalled. Moreover, the contents of the thesis were outlined.

- Chapter 2 has presented the aircraft simulation model. The equations of motion of the 6 DoF rigid body aircraft were obtained. The subsystems completing the overall simulation model were described, in particular wind gust disturbances and input–output measurement errors were taken into account. Finally, the simplified aircraft models exploited to design the residual generators, the so–called FDI models, were introduced.

- Chapter 3 has presented the PM FDI scheme. The residual generators were designed from the input–output description of the linearised aircraft model and the disturbance decoupling was obtained by computing a basis for the left null space of the disturbance distribution matrix. The residual generators design was performed in order to achieve both maximisation of a suitable fault sensitivity function and desired transient properties in terms of a fault to residual reference transfer function. Finally, the residual generators were organised into a bank structure in order to achieve fault isolation properties.

- Chapter 4 has presented the NLGA FDI scheme. The residual generators design scheme, based on the structural decoupling of the disturbance obtained by means of a coordinate transformation in the state space and in the output space, was proposed. The developed theory was applied to a simplified input affine model of

the aircraft and the residual generators for the input sensors FDI were obtained. The NLGA robustness was improved by means of a procedure based on the mixed $\mathcal{H}_-/\mathcal{H}_\infty$ optimisation of the tradeoff between fault sensitivity, disturbances and modelling. The NLGA scheme was modified in order to obtain an adaptive filters scheme, *i.e.* the NLGA–AF. In particular, the least-squares algorithm with forgetting factor was used to develop the adaptive nonlinear filters providing both the input sensors FDI and the estimation of the fault size. By combining the particle filtering algorithm with the NLGA coordinate transformation, the NLGA–PF was proposed. In particular, the basic particle filter theory was applied to obtain a particle filter for throttle sensor FDI.

- Chapter 5 has presented the simulation results. The threshold evaluation logic and the FDI procedure for a complete aircraft trajectory were described. The suggested design strategies were tested by considering a flight condition characterised by tight–coupled longitudinal and lateral dynamics. A typical aircraft reference trajectory embedding several steady–state flight conditions, such as straight phases and coordinated turns, was exploited in order to evaluate the robustness properties of the proposed PM and NLGA. A comparison with widely used data–driven and model–based FDI scheme with disturbance decoupling, such as NN and UIKF diagnosis methods, was also provided. Finally, the reliability and the robustness properties of the designed residual generators to model uncertainty, disturbances and measurements noise for the aircraft nonlinear model were investigated via Monte–Carlo simulations.

Some noticeable characteristics of FDI techniques developed in this thesis are recalled in the following:

- Concerning the PM, an important aspect is the simplicity of the technique used to generate the residuals when compared with different schemes. The algorithmic simplicity is a very important aspect when considering the need for verification and validation of a demonstrable scheme for air–worthiness certification. The more complex the computations required to implement the scheme, the higher the cost and complexity in terms of certification.

- Concerning the NLGA, the main advantage is represented by the fact that the model nonlinearities are directly taken into account. As it was shown, this fact leads to better performances in terms of fault detection promptness, with respect to other schemes.

- Concerning the NLGA–AF, in addition to a proper detection and isolation, fault size estimation is also provided. This feature is not usual for a FDI method and can be very useful during an on–line automatic flight control system reconfiguration, in order to recover a faulty operating condition. Compared with similar methods

proposed in the literature, the NLGA–AF described here has the advantage of being applicable to more general classes of nonlinear systems and less sensitive to measurement noise, since it does not use input/output signal derivatives.

- Concerning the NLGA–PF, the knowledge regarding the noise process acting on the system under diagnosis is exploited. Hence the proposed scheme provides a possible solution to nonlinear system FDI with non-Gaussian noise and disturbance.

As final remark it is worth noting that, the FDI schemes proposed in this work are of a general nature and are applicable, not only to the particular system treated in this thesis, *i.e.* the PIPER PA–30 aircraft, but to a wide class of nonlinear dynamic systems.

Appendix A

Review of Model–Based FDI

A.1 Basic Definitions

A fault is to be understood as an expected change of system function, although it may not represent physical failure or breakdown. Such a fault or malfunction hampers or disturbs the normal operation of an automatic system, thus causing an unacceptable deterioration of the performance of the system or even leading to dangerous situation. The term fault is used rather than the term failure to denote a malfunction rather that a catastrophe. The term failure suggests complete breakdown of a system component or function, whilst the term fault may be used to indicate that a malfunction may be tolerable at its present stage. A fault must be diagnosed as early as possible even it is tolerable at its early stage, to prevent any serious consequences.

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

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

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

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

Since fault identification may not be essential (if no reconfiguration action is involved), the fault diagnosis is very often considered as fault detection and isolation, abbreviated as FDI, in the literature.

FDI can be achieved using a replication of hardware (*e.g.* computers, sensors, actuators and others components) in what is known as hardware redundancy in which outputs from identical components are compared for consistency. Alternatively, FDI can be carried out using analytical or functional information about the system being monitored, *i.e.* based

on a mathematical model of the system. The latter approach is known as analytical redundancy, which is also known invariably as quantitative or model–based FDI.

The model–based FDI can be defined also as the detection, isolation and identification of faults on a system by means of methods which extract features from measured signals and use a priori information on the process available in term of a mathematical models.

Faults are thus detected by setting fixed or variable thresholds on residual signals generated from the difference between actual measurements and their estimates obtained by using the process model.

A number of residuals can be designed with each having sensitivity to individual faults occurring in different locations of the system. The analysis of each residual, once the threshold is exceeded, then leads to fault isolation.

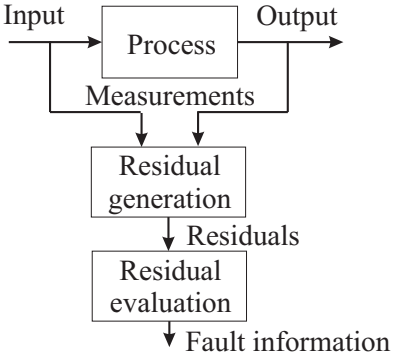


Figure A.1: Structure of model–based FDI system.

Figure A.1 shows the general and logic block diagram of model–based FDI system. It comprises two main stages of residual generation and residual evaluation. This structure was first suggested by Chow and Willsky in (Chow and Willsky 1980) and now is widely accepted by the fault diagnosis community. The two main blocks are described as follows:

Residual generation. This block generates residual signals using available inputs and outputs from the monitored system. This residual (or fault symptom) should indicate that a fault has occurred. It should normally be zero or close to zero under no fault condition, whilst distinguishably different from zero when a fault occurs. This means that the residual is characteristically independent of process inputs and outputs, in ideal conditions. Referring to Figure A.1, this block is called residual generation.

Residual evaluation. This block examines residuals for the likelihood of faults and a decision rule is then applied to determine if any faults have occurred. The residual evaluation block, shown in Figure A.1, may perform a simple threshold test (geometrical methods) on the instantaneous values or moving averages of the residuals. On the other hand, it may consist of statistical methods, *e.g.*, generalised likelihood

ratio testing or sequential probability ratio testing (Isermann 1997, Willsky 1976, Basseville 1988, Patton *et al.* 2000).

Most contributions in the field of quantitative model–based FDI focus on the residual generation problem, since the decision–making problem can be considered relatively straightforward if residuals are well–designed.

A.2 Modelling of Faulty Systems

The first step in model–based FDI approach consists of providing a mathematical description of the system under investigation which shows all the possible fault cases, as well.

The detailed scheme for FDI techniques here presented is depicted by Figure A.2. The main component are the plant under investigation, the actuators and sensors, which can be further sub–divided as input and output sensors, and finally the controller.

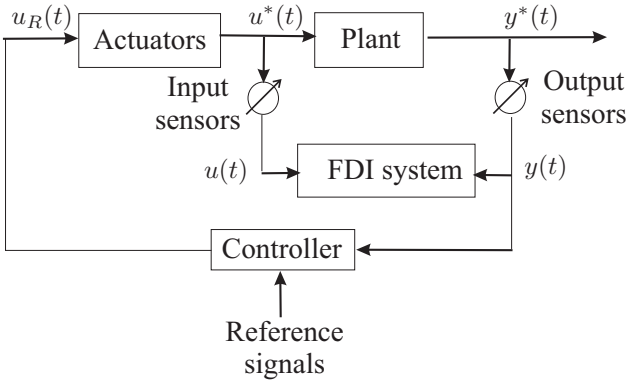


Figure A.2: Fault diagnosis in a closed–loop system.

The FDI technique presented here considers open–loop system model. In fact, as showed in Figure A.1, the information required by the FDI system is related to the open–loop system even if the system is inserted in the control loop. Hence it is not necessary to consider the controller in the design of a FDI scheme. The open–loop subsystem considered for the FDI design is illustrated in Figure A.3, where also the faults acting on the various subsystems are represented.

Under the hypothesis of linearity, process dynamics can be described by the following continuos–time, time–invariant, linear dynamic system in the state–space form

$$\begin{aligned} \dot{x}(t) &= A\,x(t) + B\,u^*(t) \\ y^*(t) &= C\,x(t) \end{aligned} \tag{A.1}$$

where $x(t) \in \Re^n$ is the system state vector, $u^*(t) \in \Re^r$ is the input signal vector driven by actuators, and $y^*(t) \in \Re^m$ is the real system output vector. A, B, and C are system matrices with appropriate dimensions.

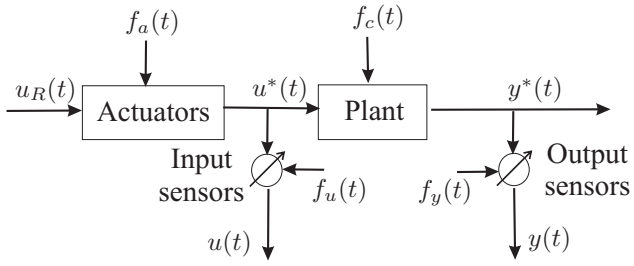


Figure A.3: The monitored system and fault topology.

With reference to Figure A.3, a component fault vector $f_c(t)$ affects process dynamics as follows

$$\dot{x}(t) = A\,x(t) + B\,u^*(t) + f_c(t) \tag{A.2}$$

In some cases, component faults come from a change in the system parameters, *e.g.* a change in entries of the A matrix. For example, a change in the i-th row and the j-th column of the A matrix, leads to a fault vector $f_c(t)$ described as

$$f_c(t) = I_n^i\,\Delta a_{ij}\,x_j(t) \tag{A.3}$$

where $x_j(t)$ in the j-th element of the vector $x(t)$.

As the actual process output $y^*(t)$ is not directly available, a sensor is used to acquire a measure of the system outputs. Moreover, generally speaking, a sensor can be also used to measure the system inputs $u^*(t)$ (*e.g.* for uncontrolled system). By neglecting sensor dynamics, faults on input and output sensors are modelled with additive signals, respectively, as

$$\begin{aligned} u(t) &= u^*(t) + f_u(t) \\ y(t) &= y^*(t) + f_y(t) \end{aligned} \tag{A.4}$$

where the vectors $f_u(t) = [f_{u_1}(t) \dots f_{u_r}(t)]^T$ and $f_y(t) = [f_{y_1}(t) \dots f_{y_m}(t)]^T$ are chosen to describe a fault situation. Usually fault modes can be described by step and ramp signals in order to model abrupt and incipient (hard to detect) faults, representing bias and drift, respectively.

Moreover, for technical reasons, sensor output signals are generally affected by measurement noise. In this case (A.4) has to be replaced by

$$\begin{aligned} u(t) &= u^*(t) + \tilde{u}(t) + f_u(t) \\ y(t) &= y^*(t) + \tilde{y}(t) + f_y(t) \end{aligned} \tag{A.5}$$

in which the sequences $\tilde{u}(t)$ and $\tilde{y}(t)$ are usually described as white, zero–mean, uncorrelated Gaussian processes.

With reference to a controlled system, according to Figure A.3, signals $u^*(t)$ are the actuator response to the command signals $u_R(t)$. A purely algebraic actuator (*i.e.* with

gain equal to 1) can be described by

$$u^*(t) = u_R(t) + f_a(t) \tag{A.6}$$

where, similarly to input-output sensor fault situation, $f_a(t) \in \Re^r$ is the actuator fault vector.

In general, as shown in Figure A.3, if the actuation signals $u^*(t)$ are assumed to be measurable, by neglecting input and output sensor noises, the process model with fault can be described by the following system equation

$$\begin{aligned} \dot{x}(t) &= A\,x(t) + f_c(t) - B\,f_u(t) + B\,u(t) \\ y(t) &= Cx(t) + f_y(t) \end{aligned} \tag{A.7}$$

On the other hand, Figure A.4 represents the case where the $u_R(t)$ signals are measured only by the input sensors. Such a configuration represents a critical situation with respect to the input sensor connection depicted in Figure A.3.

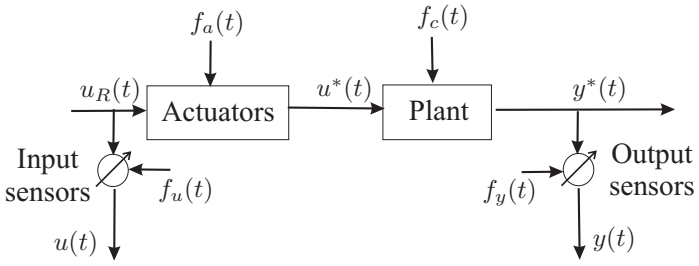


Figure A.4: Fault topology with actuator input signal measurement.

In this situation, actuator faults cannot be directly related to the input measurements $u(t)$ but their effects can only be detected by means of output signals $y(t)$. By taking into account also actuator faults $f_a(t)$, the description below is obtained

$$\begin{aligned} \dot{x}(t) &= A\,x(t) + f_c(t) + B\,f_a(t) - B\,f_u(t) + B\,u(t) \\ y(t) &= C\,x(t) + f_y(t) \end{aligned} \tag{A.8}$$

Moreover, considering the general case, a system affected by all possible faults can be described by the the following state–space model

$$\begin{aligned} \dot{x}(t) &= A\,x(t) + B\,u(t) + L_1 f(t) \\ y(t) &= C\,x(t) + L_2\,f(t) \end{aligned} \tag{A.9}$$

where entries of the vector $f(t) = [f_a^{\mathrm{T}},\ f_u^{\mathrm{T}},\ f_c^{\mathrm{T}},\ f_y^{\mathrm{T}}]^{\mathrm{T}}$ correspond to specific faults. In practice, it is reasonable to assume that the fault signals are described by unknown time functions. The matrices L_1, L_2 are known as faulty entry matrices which describe how

the faults enter the system. The vectors $u(t)$ and $y(t)$ are the available and measurable inputs and outputs, respectively. Both vectors are supposed known for FDI purpose.

The state–space model (A.9) can be expressed also in the input–output transfer matrix representation

$$y(t) = G_{yu}(s)\, u(t) + G_{yf}(s)\, f(t) \tag{A.10}$$

where

$$\begin{aligned} G_{yu}(s) &= C\,(s\,I - A)^{-1}\,B \\ G_{yf}(s) &= C\,(s\,I - A)^{-1}\,L_1 + L_2 \end{aligned} \tag{A.11}$$

Both the general models for FDI described by (A.9) and (A.10) in the time and frequency domain, respectively, have been widely accepted in the fault diagnosis literature (Patton *et al.* 1989, Patton *et al.* 2000, Chen and Patton 1999, Gertler 1998).

Under these assumptions, the general model–based FDI problem here treated can be performed on the basis of the knowledge only of the measured sequences $u(t)$ and $y(t)$.

A.3 Residual Generator General Structure

The most frequently used FDI methods exploit the a priori knowledge of characteristics of certain signals. As an example, the spectrum, the dynamic range of the signal and its variations may be checked. However, the necessity of a priori information concerning the monitored signals and the dependence of the signal characteristics on unknown working conditions of the system under diagnosis are main drawbacks of such a class of methods.

The most significant contribution in modern model–based approaches is the introduction of the symptom or residual signals, which depend on faults and are independent of system operating states. They represent the inconsistency between the actual system measurements and the corresponding signals of the mathematical model.

The residual generator block introduced in Figure A.1 can be interpreted as illustrated in Figure A.5 (Basseville 1988). The auxiliary redundant signal $z(t)$ is generated by the function $W_1\big(u(\cdot), y(\cdot)\big)$ and, together with the measurement $y(t)$, the symptom signal $r(t)$ is computed by means of $W_2\big(z(\cdot), y(\cdot)\big)$. In the fault–free case, the following relations are satisfied

$$\begin{aligned} z(t) &= W_1\big(u(\cdot), y(\cdot)\big) \\ r(t) &= W_2\big(z(\cdot), y(\cdot)\big) = 0 \end{aligned} \tag{A.12}$$

and, when a fault occurs in the plant, the residual $r(t)$ will be different from zero.

The simplest residual generator is obtained when the system W_1 is a plant identical model $z(t) = W_1\big(u(\cdot)\big)$. The measurement $y(t)$ is not required in W_1 because it is a system simulator. The signal $z(t)$ represents the simulated output and the residual is computed as $r(t) = z(t) - y(t)$. Since it is an open–loop system, the process simulation may become unstable.

An extension to the simplest residual generator is to replace $W_1\big(u(\cdot)\big)$ by $W_1\big(u(\cdot), y(\cdot)\big)$, *i.e.* an output estimator fed by both system input and output. In such a case, function W_1 generates an estimation of a linear function of the output $W_1\big(u(\cdot), y(\cdot)\big) = My(t)$ whilst

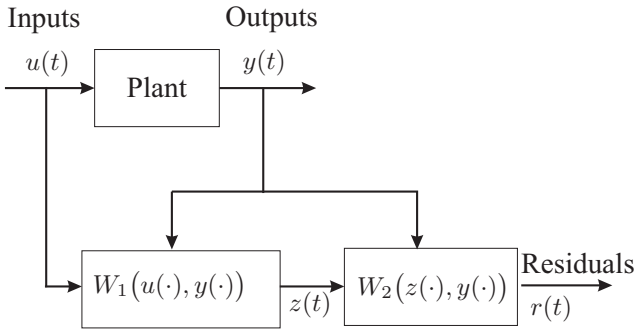


Figure A.5: Residual generator general structure.

function W_2 can be defined as $W_2\big(z(\cdot), y(\cdot)\big) = W\big(z(t) - My(t)\big)$, W being a weighting matrix.

Figure A.6 represents a general structure for all residual generators using the input–output transfer matrix description. With reference to (A.10) and (A.11), the residual generator structure is expressed mathematically by the generalised representation

$$r(t) = \begin{bmatrix} H_u(s) & H_y(s) \end{bmatrix} \begin{bmatrix} u(t) \\ y(t) \end{bmatrix} = H_u(s)\, u(t) + H_y(s)\, y(t) \tag{A.13}$$

where $H_u(s)$ and $H_y(s)$ are transfer matrices which can be designed using stable linear systems.

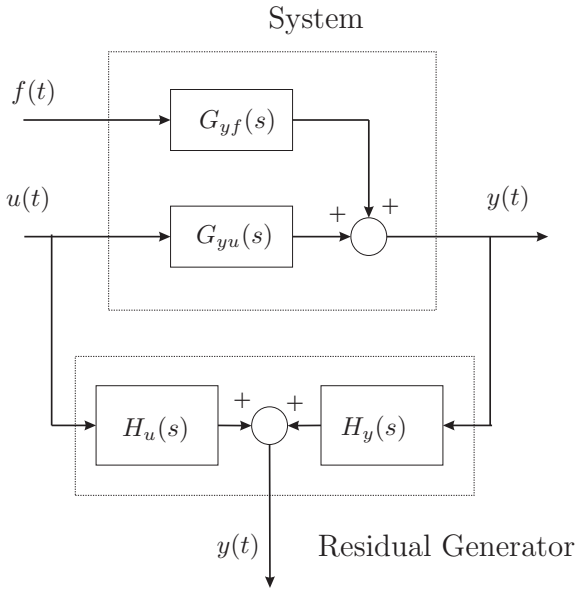


Figure A.6: Residual generator using input–output description.

According to the definition, the residual $r(t)$ has to be designed to become zero for fault–free case and different from zero in case of failures. This means that

$$r(t) = 0 \quad \text{if and only if} \quad f(t) = 0 \tag{A.14}$$

In order to satisfy (A.14), the design of the transfer matrices $H_u(s)$ and $H_y(s)$ must satisfy to the constraint conditions

$$H_u(s) + H_y(s)\, G_{yu} = 0 \tag{A.15}$$

It is worth noting that different residual generators can be obtained by using different parametrisations of $H_u(s)$ and $H_y(s)$ (Patton and Chen 1991*a*).

A.4 Fault Detectability and Isolability

By substituting (A.10) and (A.15) into (A.13), it results that, when faults occur in the monitored system, the response of the residual vector is

$$r(t) = H_y(s)\, G_{yf}(s)\, f(t) \tag{A.16}$$

To detect the i–th fault in the residual $r(t)$, the i–th column $[H_y(s)\, G_{yf}(s)]_i$ of the transfer matrix $[H_y(s)\, G_{yf}(s)]$ should be nonzero, especially for steady values, *i.e.*

$$[H_y(s)\, G_{yf}(s)]_i \neq 0 \quad \text{and} \quad [H_y(0)\, G_{yf}(0)]_i \neq 0 \tag{A.17}$$

If those conditions are satisfied, the i–th fault can be said to be detectable using the residual.

The isolability of faults is also an important issue. In precise terms, isolability is the ability of the procedure to distinguish (isolate) certain specific faults. Although a single residual signal is sufficient to detect faults, a set of residuals (or a residual vector) is required for fault isolation. To facilitate the isolation problem, residual sets are usually generated in one of the following ways:

Structured residual set: to generate a set of residuals, each residual is designed to be sensitive to a subset of faults while remaining insensitive to other faults (Chow and Willsky 1984, Gertler 1998).

Directional residual vectors: to make the residual vector lie in fixed and fault–specific directions (or subspaces) in the residual space in response to a specific fault (Gertler 1998, Beard 1971).

The basis for the isolation of a fault is the fault signature, *i.e.* a feature obtained from a diagnostic model defining the effects associated with a fault (Beard 1971, Jones 1973). Each signature must be uniquely related to one fault to accomplish fault isolation.

In presence of faults on the input–output sensors of the monitored system the input–output relation (A.10) can be rewritten as follows

$$y(t) = G_{yu}(s)\, u(t) + G_{yf_u}(s)\, f_u(t) + f_s(t) \tag{A.18}$$

where the transfer matrix $G_{yf_u}(s)$ represents the effect on the system of the input sensor faults.

When output sensor faults occur, the residual vector becomes

$$r(t) = H_y(s)\, f_s(t) \tag{A.19}$$

Here the fault transfer matrix $H_y(s)$ will be chosen according to specific requirements. For the isolation purpose, we can make $r(t)$ independent to the i–th sensor fault by simply making the i–th column of $H_y(s)$ equal to zero. If $H_y(s)$ is a diagonal matrix, each element of the residual is only affected by a specified sensor fault; this is very useful for fault isolation. The only constraint on $H_y(s)$ is that it be stable and implementable. Once it has been chosen, $H_u(s)$ can be determined by the constraint (A.15).

When input sensor faults occur, the residual vector becomes

$$r(t) = H_y(s)\, G_{yf_u}(s)\, f_s(t) \tag{A.20}$$

In this case, the fault transfer function matrix is $H_y(s)\, G_{yf_u}(s)$. To make $r(t)$ independent of the i–th sensor fault, the choice of $H_y(s)$ must ensure that the i–th column of matrix $H_y(s)\, G_{yf_u}(s)$ is equal to zero. For any given fault transfer matrix $H_y(s)\, G_{yf_u}(s)$, the stable and implementable $H_y(s)$ does not always exist. That is to say, we do not have full freedom to achieve required input sensor fault isolation.

A.5 Residual Generation Techniques

The generation of symptoms is the main issue in model–based fault diagnosis. A variety of methods are available in literature for residual generation and the most common are briefly summarised in the following:

Observer–based approaches. The basic idea within this approach is to estimate the outputs of the system from the measurements or a subset of measurements by using either Luenberger observers in a deterministic setting or Kalman filters in stochastic setting. Then, the output estimation error or innovations in the stochastic case, is used as a residual. The flexibility in selecting observer gains has been fully exploited in the literature yielding a rich variety of fault detection schemes (Beard 1971, Frank 1993, Frank and Ding 1997, Patton and Chen 1997, Willsky 1976, Basseville 1988).

Parity vector (relation) methods. This approach is based on checking the consistency of the mathematical relations between the outputs (or a subset of outputs) and inputs. These relations may lead to direct redundancy, which gives the static

algebraic relations between the sensor outputs or they may be lead to temporal redundancy, which gives the dynamic relations between inputs and outputs (Chow and Willsky 1984, Gertler and Singer 1990, Patton and Chen 1991*a*, Gertler and Monajemy 1993, Delmaire *et al.* 1999).

Fault detection via parameter estimation. This approach is based on the assumption that the faults are reflected in the physical system parameters and the basic idea is that the parameters of the actual process are estimated on–line using well–known parameter estimations methods. The results are thus compared with the parameters of the reference model; obtained initially under fault–free assumptions. Any discrepancy can indicate that a fault may have occurred (Isermann 1984, Isermann and Freyermuth 1992, Isermann 1993, Isermann and Ballé 1997, Patton *et al.* 2000).

Though conceptually different, it has been shown that there are close relationships among these approaches. It is easy to see that the parity space approach leads to a parallel model which can be interpreted as a special class of observer, namely the so–called dead–beat observer with all poles at the origin (Massoumnia 1986, Frank 1990). This means that the residual generator resulting from the parity equation approach can be subsumed, as a special case, under the group of diagnostic observers.

A.6 Residual Evaluation

When the residual generation stage has been performed, the second step requires the examination of symptoms in order to determine if any faults have occurred.

The simplest and most widely used way to fault detection is achieved by directly comparing residual signal $r(t)$ or a residual function $J(r(t))$ with a fixed threshold ϵ or a threshold function $\varepsilon(t)$ as follows

$$\begin{array}{lll} J(r(t)) \leq \varepsilon(t) & \text{for} & f(t) = 0 \\ J(r(t)) > \varepsilon(t) & \text{for} & f(t) \neq 0 \end{array} \tag{A.21}$$

where $f(t)$ is the general fault vector defined in (A.9). If the residual exceeds the threshold, a fault may be occurred. This test works especially well with fixed thresholds ε if the process operates approximately in a steady state and it reacts after relatively large feature, *i.e.* after either a large sudden or a long-lasting gradually increasing fault. On the other hand, adaptive thresholds $\varepsilon(t)$ can be exploited which depend on plant operating conditions, for example when $\varepsilon(t)$ is expressed as a function of plant inputs (Clark 1989, Chen and Patton 1999).

Because of the presence of noise, disturbances and other unknown signals acting upon the monitored system, the decision making process can exploits statistical methods. In this case, the measured or estimated quantities, such as signals, parameters, state variables

or residuals are usually represented by stochastic variables $r(t) = \{r_i(t)\}_i^q$, with mean value and variance (Willsky 1976)

$$\bar{r}_i = E\{r_i(t)\} \qquad \bar{\sigma}_i^2 = E\{[r_i(t) - \bar{r}_i]^2\} \tag{A.22}$$

as normal values for the fault-free process. Analytic symptoms are then obtained as changes

$$\Delta r_i = E\{r_i(t) - \bar{r}_i\} \qquad \Delta\sigma_i = E\{\sigma_i(t) - \bar{\sigma}_i\} \tag{A.23}$$

with reference to the normal values. In order to separate normal from faulty behaviour, usually a fixed threshold Δr_{tol} defined as

$$\Delta r_{tol} = \epsilon\bar{\sigma}_r \qquad \epsilon \geq 2 \tag{A.24}$$

has to be selected. By a proper choice of ϵ, a compromise has to be made between the detection of small faults and false alarms.

Another class of methods can be exploited for detecting residual changes due to faults. Therefore, techniques of change detection, *e.g.* as a likelihood–ratio–test or Bayes decision, a run–sum test are commonly used (Isermann 1984, Basseville and Benveniste 1986, Basseville and Nikiforov 1993). Moreover, fuzzy or adaptive thresholds may improve the binary decision (Chen and Patton 1999, Patton *et al.* 2000).

A.7 Robustness Problem

The model–based FDI uses a mathematical system model. The better the model represents the system, the better will be the reliability and performance in FDI. However, modelling errors and disturbances are inevitable, and hence there is a need to develop robust FDI algorithm.

A robust FDI system is sensitive only to faults, even in the presence of a model–reality mismatch. To achieve robustness in FDI, the residual should be insensitive to uncertainty, whilst sensitive to faults, and therefore robust (Chow and Willsky 1984, Ding and Frank 1990, Frank 1994, Frank and Ding 1997, Patton and Chen 1994).

In order to summarise the approach to the robustness problem, the state–space model of the monitored system should be considered (Patton and Chen 1993)

$$\begin{aligned} \dot{x}(t) &= (A + \Delta A)\, x(t) + (B + \Delta B)\, u(t) + E_1\, d(t) + L_1\, f(t) \\ y(t) &= C\, x(t) + E_2\, d(t) + L_2\, f(t) \end{aligned} \tag{A.25}$$

where $d(t)$ is the disturbance vector; E_1 and E_2 are the known input distribution matrices; the matrices ΔA and ΔB are the parameter errors or variations which represent modelling errors.

The transfer matrix description between the output $y(t)$ and input $u(t)$ of the system (A.25) is then

$$y(t) = (G_{yu}(s) + \Delta G_{yu}(s))\, u(t) + G_{yd}(s)\, d(t) + G_{yf}\, f(t) \tag{A.26}$$

where $\Delta G_{yu}(s)$ is used to describe modelling errors, whilst $G_{yd}(s)$ describes the disturbance.

With reference to the residual generator of Figure A.6 and described by (A.13), the residual vector has to be rewritten as

$$r(t) = H_y\, G_{yf}(s)\, f(t) + H_y\, \Delta G_{yu}(s) + H_y(s)\, G_{yd}\, d(t) \tag{A.27}$$

With respect to (A.13), the terms $H_y(s)\, G_{yd}(s)$ and $H_y(s)\, \Delta G_{yu}(s)$ cannot be deleted. Both faults and modelling uncertainty (disturbance and modelling error) affect the residual and hence discrimination between these two effects is difficult.

The principle of disturbance de–coupling for robust residual generation requires that the residual generator satisfies

$$H_y(s)\, G_{yd}(s) = 0 \tag{A.28}$$

in order to achieve total de–coupling between residual $r(t)$ and disturbance $d(t)$. This property can be achieved by using the unknown input observer (Watanabe and Himmelblau 1982, Wünnenberg and Frank 1987, Chen *et al.* 1996, Frank *et al.* 2000), optimal (robust) parity relations (Chow and Willsky 1984, Lou *et al.* 1986, Wünnenberg 1990, Wünnenberg and Frank 1990, Frank *et al.* 2000) or alternatively the eigenstructure assignment approach (Patton *et al.* 1986, Patton and Chen 1991*b*, Liu and Patton 1998, Patton and Chen 2000, Duan *et al.* 2002).

If condition (A.28) does not hold, perfect (accurate) de–coupling is not achievable. One can consider an optimal or approximate de–coupling by minimising the following performance index over a specified frequency range (Frank *et al.* 2000)

$$J = \frac{\| H_y(j\omega)\, G_{yd}(j\omega) \|}{\| H_y(j\omega)\, G_{yf}(j\omega) \|} \tag{A.29}$$

On the other hand, with reference to the modelling errors in (A.27), represented by the term $\Delta G_{yu}(z)$ the robust problem is more difficult to solve. Two main techniques have been described by Patton and Chen. In the first case, the uncertainty is taken into account at the residual design stage (Chen *et al.* 1996); this is known as active robustness in fault diagnosis (Patton and Chen 1994).

The active way of achieving a robust solution is to approximate uncertainties, *i.e.* representing approximately modelling errors as disturbances (Chen and Patton 1999)

$$\Delta G_{yu}(s)\, u(t) \approx G_{yd_1}(s)\, d_1(t) \tag{A.30}$$

where $d_1(t)$ is an unknown vector and $G_{yd_1}(s)$ is an estimated transfer function. When this approximate structure is exploited to design disturbance de–coupling residual generators, robust FDI can be achieved.

The second approach called passive robustness makes use of a residual evaluator with adaptive threshold. As a simple example, it is assumed that the residual generation uncertainty (A.27) is only represented by modelling errors. The fault–free residual $r(t)$ is

$$r(t) = H_y(s)\, \Delta G_{yu}(s)\, u(t) \tag{A.31}$$

Under the assumption that the modelling errors are bounded by a value δ, such that

$$\| \Delta G_{yu}(w) \| \leq \delta \tag{A.32}$$

an adaptive threshold $\varepsilon(t)$ can be generated by a linear system

$$\varepsilon(t) = \delta \, H_y(s) \, u(t) \tag{A.33}$$

In such case, the threshold $\varepsilon(t)$ is no longer fixed but depend on the input $u(t)$, thus being adaptive to the system operating point. A fault is then detected if

$$\| r(t) \| > \| \varepsilon(t) \| \tag{A.34}$$

A robust FDI technique with the threshold adaptor or selector is therefore briefly recalled (Clark 1989, Emami-Naeini *et al.* 1988, Ding and Frank 1991). This method represents a passive approach since no effort is made to design a robust residual.

Bibliography

Amato, Francesco, Carlo Cosentino, Massimiliano Mattei and Gaetano Paviglianiti (2006). A direct/functional redundancy scheme for fault detection and isolation on an aircraft. *Aerospace Science and Technology* **10**(4), 338–345.

ARINC (1998). Aeronautical Radio Inc.. In: *MD. ARINC characteristic 706–4, Mark 5 subsonic air data system*. Annapolis.

Basseville, M. (1988). Detecting changes in signals and systems: A survey. *Automatica* **24**(3), 309–326.

Basseville, M. and A. Benveniste (1986). Detection of abrupt changes in signals and and dynamical systems. In: *Lecture Notes in Control and Information Sciences*. Vol. 77. Springer–Verlag. London.

Basseville, M. and I. V. Nikiforov (1993). *Detection of Abrupt Changes: Theory and Application*. Prentice–Hall Inc.

Beard, R. V. (1971). Failure accomodation in linear systems through self-reorganisation. Technical Report MVT-71-1. Man Vehicle Lab.. Cambridge, Mass.

Beghelli, S., M. Benini and S. Simani (2007*a*). Residual Generator Design for the FDI of Linear Multivariable Dynamic Systems. In: *European Control Conference 2007 – ECC'07* (ICCS EUCA, Ed.). Vol. CD–Rom. EUCA, ICCS, IFAC, ACPA & IEEE CSS. Kos, Greece. pp. 2288–2295.

Beghelli, S., M. Benini, G. Bertoni, M. Bonfè, P. Castaldi, W. Geri and S. Simani (2007*b*). Design of robust fault diagnosis schemes for a simulated aircraft nonlinear model. In: *5th Workshop on Advanced Control and Diagnosis – ACD2007*. Vol. CD–Rom. IAR – Institute for Automation and Robotics, ICD Working Group. Grenoble, France. pp. 1–6. Organized by IAR ICD Working Group.

Benini, M., M. Bonfè, P. Castaldi and S. Simani (2009). Nonlinear Geometric Approach–Based Filtering Methods for Aircraft Actuator FDI. In: *7th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes – SAFEPROCESS 2009*. Iternational Federation of Automatic Control IFAC – Advanced Control Systems SAC – Universitat Politècnica de Catalunya UPC. Barcelona, Spain. accepted.

Benini, M., M. Bonfè, P. Castaldi, W. Geri and S. Simani (2008*a*). Design and Performance Evaluation of Fault Diagnosis Strategies for a Simulated Aircraft Nonlinear Model. *Journal of Control Science and Engineering* **2008**, 1–18. Special Issue on "Robustness Issues in Fault Diagnosis and Fault Tolerant Control". Published by Hindawi Publishing Corporation. ISSN (printed): 1687-5249. ISSN (electronic): 1687-5257.

Benini, M., M. Bonfè, P. Castaldi, W. Geri and S. Simani (2008*b*). Fault Diagnosis Strategies for a Simulated Nonlinear Aircraft Model. In: *Proceedings of the 17th IFAC World Congress* (Myung Jin Chung, Pradeep Misra and Hyungbo Shim, Eds.). Vol. CD–Rom. The International Federation of Automatic Control (IFAC). IFAC. Seoul, Korea. pp. 7300–7307. Paper Id: WeA01.19.

Bonfè, M., P. Castaldi, W. Geri and S. Simani (2007*a*). Design and Performance Evaluation of Residual Generators for the FDI of an Aircraft. *International Journal of Automation and Computing* **4**(2), 156–163. DOI: 10.1007/s11633–007–0156–7.

Bonfè, M., P. Castaldi, W. Geri and S. Simani (2007*b*). Nonlinear Actuator Fault Detection and Isolation for a General Aviation Aircraft. *Space Technology – Space Engineering, Telecommunication, Systems Engineering and Control* **27**(2–3), 107–113. Special Issue on Automatic Control in Aerospace.

Bonfè, M., P. Castaldi, W. Geri, S. Simani and M. Benini (2008). Fault Diagnosis Techniques for Aircraft Simulated Model Sensors. In: *23rd IAR Workshop on Advanced Control and Diagnosis – IAR/ACD2008* (K. J. Burnham and J. G. Linden, Eds.). Vol. CD–Rom. IAR – Institute for Automation and Robotics, CTAC Coventry University. Coventry University, Coventry, UK. pp. 96–101. Organized by IAR CTAC Working Group.

Bonfè, M., S. Simani, P. Castaldi and W. Geri (2004). Residual generator computation for fault detection of a general aviation aircraft. In: *ACA 2004. 16^{th} IFAC Symposium on Automatic Control in Aerospace*. Vol. 2. IFAC. IFAC. St. Petersburg, Russia. pp. 318–323.

Bonfè, Marcello, Paolo Castaldi, Walter Geri and Silvio Simani (2006). Fault Detection and Isolation for On–Board Sensors of a General Aviation Aircraft. *International Journal of Adaptive Control and Signal Processing* **20**(8), 381–408. Copyright 2006 John Wiley & Sons, Ltd. ISSN: 0890-6327.

Bryson, A. E. Jr. (1994). *Control of Spacecraft and Aircraft*. Princeton University, UK. UK. ISBN 0-691-08782-2.

Castaldi, P., W. Geri, M. Bonfè, S. Simani and M. Benini (2009). Design of residual generators and adaptive filters for the fdi of aircraft model sensors. *Control Engineering Practice*. ACA'07 – 17th IFAC Symposium on Automatic Control in Aerospace Special Issue. Publisher: Elsevier Science. ISSN: 0967–0661. doi:10.1016/j.conengprac.2008.11.006.

Chen, J. and R. J. Patton (1999). *Robust Model–Based Fault Diagnosis for Dynamic Systems*. Kluwer Academic Publishers.

Chen, J., R. J. Patton and H. Y. Zhang (1996). Design of unknown input observer and robust fault detection filters. *Int. J. Control* **63**(1), 85–105.

Chow, E. Y. and A. S. Willsky (1980). Issue in the development of a general algorithm for reliable failure detection. In: *Proc. of the 19th Conf. on Decision & Control*. Albuquerque, NM.

Chow, E. Y. and A. S. Willsky (1984). Analytical redundancy and the design of robust detection systems. *IEEE Trans. Automatic Control* **29**(7), 603–614.

Clark, R. N. (1989). *Fault Diagnosis in Dynamic Systems: Theory and Application*. Chap. 2, pp. 21–45. Prentice Hall.

Collinson, R. P. G. (2002). *Introduction to Avionics Systems*. 2nd ed.. Springer–Verlag. ISBN: 1402072783.

De Persis, C. and A. Isidori (2000). On the observability codistributions of a nonlinear system. In: *Systems and Control Letters*. Vol. 40. pp. 297–304.

De Persis, C. and A Isidori (2001). A geometric approach to non–linear fault detection and isolation. *IEEE Transactions on Automatic Control* **45**(6), 853–865.

De Persis, C., R. De Sanctis and A. Isidori (2001). Nonlinear actuator fault detection and isolation for a VTOL aircraft. *Proceedings of the American Control Conference* pp. 4449–4454.

Delmaire, G., Ph. Cassar, M. Staroswiecki and C. Christophe (1999). Comparison of multivariable identification and parity space techniques for FDI purpose in M.I.M.O. systems. In: *ECC'99*. Karlsruhe, Germany.

Ding, Steven X. (2008). *Model–based Fault Diagnosis Techniques: Design Schemes, Algorithms, and Tools*. 1st ed.. Springer. Berlin Heidelberg. ISBN: 978–3540763031.

Ding, X. and P. M. Frank (1990). Fault detection via factorization approach. *Syst. Contr. Lett.* **14**(5), 431–436.

Ding, X. and P. M. Frank (1991). Frequency domain approach and threshold selector for robust model–based fault detection and isolation. In: *Preprint of IFAC/IMACS Symposium SAFEPROCESS'91*. Vol. 1. pp. 307–312. Baden–Baden.

Doucet, A., de Freitas, N. and Gordon, N., Eds. (2001). *Sequential Monte Carlo Methods in Practice*. Statistics for Engineering and Information Science. Springer–Verlag. New York. ISBN: 978–0387951461.

Duan, G.R., D. How and R.J. Patton (2002). Robust Fault Detection in Descriptor Systems via Generalised Unknown Input Observers. *Int. J. Systems Science*.

Emami-Naeini, A.E., M.M. Akhter and M.M. Rock (1988). Effect of model uncertainty on failure detection: the threshold selector. *IEEE Trans. on Automatic Control* **33**(12), 1105–1115.

Etkin, B. and L. D. Reid (1996). *Dynamics of Flight* – Stability and Control. 3nd ed.. John Wiley and Sons. ISBN: 0471034185.

Fink, M.P. and D.C. Jr. Freeman (1969). Full-scale wind-tunnel investigation of static longitudinal and lateral characteristics of a light twin-engine airplane. Technical Report TN D-4983. N.A.S.A.

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

Frank, P. M. (1993). Advances in observer-based fault diagnosis. *Proc. TOOLDIAG'93 Conference*. CERT, Toulose (F).

Frank, P. M. (1994). Enhancement of robustness on observer-based fault detection. *International Journal of Control* **59**(4), 955–983.

Frank, P. M. and X. Ding (1997). Survey of robust residual generation and evaluation methods in observer-based fault detection system. *Journal of Process Control* **7**(6), 403–424.

Frank, P. M., Steven X. Ding and Birgit Köpper-Seliger (2000). Current Developments in the Theory of FDI. In: *SAFEPROCESS'00: Preprints of the IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes*. Vol. 1. Budapest, Hungary. pp. 16–27.

Frisk, Erik and M. Nyberg (2001). A minimal polynomial basis solution to residual generation for fault diagnosis in linear systems. *Automatica* **37**(9), 1417–1424.

Germani, A., C. Manes and P. Palumbo (2007). Filtering of Stochastic Nonlinear Differential Systems via a Carleman Approximation Approach. *IEEE Transactions ON Automatic Control* **52**(11), 2166–2172.

Gertler, J. (1998). *Fault Detection and Diagnosis in Engineering Systems*. Marcel Dekker. New York.

Gertler, J. and D. Singer (1990). A new structural framework for parity equation–based failure detection and isolation. *Automatica* **26**(2), 381–388.

Gertler, J. and R. Monajemy (1993). Generating directional residuals with dynamic parity equations. *Proc. of the 12th IFAC World Congress* **7**, 505–510. Sydney.

Guidorzi, R. P. (1975). Canonical Structures in the Identification. *Automatica* **11**, 361–374.

Hammouri, H., M. Kinnaert and E.H. El Yaagoubi (1999). Observer–based approach to fault detection and isolation for nonlinear systems. *IEEE Transactions on Automatic Control* **44**(10), 1879–1884.

Hou, M. and R. J. Patton (1996). An LMI approach to H_-/H_∞ fault detection observers. In: *CONTROL'96*. IEE. University of Exeter, UK. pp. 305–310.

Isermann, R. (1984). Process fault detection based on modeling and estimation methods: A survey. *Automatica* **20**(4), 387–404.

Isermann, R. (1993). Fault diagnosis via parameter estimation and knowledge processing. *Automatica* **29**(4), 815–835.

Isermann, R. (1997). Supervision, fault detection and fault diagnosis methods: an introduction. *Control Engineering Practice* **5**(5), 639–652.

Isermann, R. and Freyermuth, B., Eds. (1992). *Fault Detection, Supervision and Safety for Technical Processes*. Vol. 6 of *IFAC Symposia Series*. SAFEPROCESS'91. Pergamon Press.

Isermann, R. and P. Ballé (1997). Trends in the application of model-based fault detection and diagnosis of technical processes. *Control Engineering Practice* **5**(5), 709–719.

Isermann, Rolf (2005). *Fault–Diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance*. 1st ed.. Springer–Verlag. ISBN: 3540241124.

Jones, H. L. (1973). Failure detection in linear systems. PhD thesis. Dept. of Aeronautics, M.I.T.. Cambridge, Mass.

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

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

Kailath, T. (1980). *Linear systems*. Prentice Hall. Englewood Cliffs, New Jersey 07632.

Korbicz, J., Koscielny, J. M., Kowalczuk, Z. and Cholewa, W., Eds. (2004). *Fault Diagnosis: Models, Artificial Intelligence, Applications*. 1st ed.. Springer–Verlag. ISBN: 3540407677.

Koziol, J.S. Jr. (1971). Simulation model for the Piper PA–30 light maneuverable aircraft in the final approach. Technical Report DOT-TSC-FAA-71-11. N.A.S.A.

Liu, G. P. and R. J. Patton (1998). *Eigenstructure Assignment for Control System Design*. John Wiley & Sons. England.

Lou, X., A.S. Willsky and G.C. Verghese (1986). Optimal robust redudancy relations for failure detection in uncertainty systems. *Automatica* **22**(3), 333–344.

Marcos, Andrès, Subhabrata Ganguli and Gary J. Balas (2005). An application of H_∞ fault detection and isolation to a transport aircraft. *Control Engineering Practice* **13**(1), 105–119.

Massoumnia, M. A. (1986). A geometric appoach to failure detection and identification in linear systems. PhD thesis. Massachusetts Institute of Technology. Massachusetts, USA.

Moorhouse, D. and R. Woodcock (1980). U.S. Military Specification MIL–F–8785C. Technical report. U.S. Department of Defense.

Ojha, S. K. (1995). *Flight Performance of Aircraft*. AIAA Aerospace Educational Series. AIAA.

Patton, R. J. and J. Chen (1991*a*). A review of parity space approaches to fault diagnosis. In: *IFAC Symposium SAFEPROCESS '91*. Baden-Baden.

Patton, R. J. and J. Chen (1991*b*). Robust fault detection using eigenstructure assignment: a tutorial consideration and some new results. *30*–th *IEEE Conference on Decision and Control* pp. 2242–2247.

Patton, R. J. and J. Chen (1994). A review of parity space approaches to fault diagnosis for aerospace systems. *AIAA J. of Guidance, Contr. & Dynamics* **17**(2), 278–285.

Patton, R. J. and J. Chen (1997). Observer–based fault detection and isolation: Robustness and applications. *Control Eng. Practice* **5**(5), 671–682.

Patton, R. J. and J. Chen (2000). On eigenstructure assignment for robust fault diagnosis. *Int. J. of Robust & Non–Linear Control*.

Patton, R. J., Frank, P. M. and Clark, R. N., Eds. (1989). *Fault Diagnosis in Dynamic Systems, Theory and Application*. Control Engineering Series. Prentice Hall. London.

Patton, R. J., Frank, P. M. and Clark, R. N., Eds. (2000). *Issues of Fault Diagnosis for Dynamic Systems*. Springer–Verlag. London Limited.

Patton, R.J. and J. Chen (1993). Optimal selection of unknown input distribution matrix in the design of robust observers for fault diagnosis. *Automatica* **29**(4), 837–841.

Patton, R.J., S.W. Willcox and J.S. Winter (1986). A parameter insensitive technique for aircraft sensor fault diagnosis using eigenstructure assignment and analytical redundancy. In: *Proc. of the AIAA Conference on Guidance, Navigation & Control*. number 86–2029–CP. Williamsburg, VA.

Randle, J.S. and M.A. Horton (1997). Low cost navigation using micro–machined technology. In: *ITSC'97 – IEEE Conference on Intelligent Transportation System*. pp. 1064–1067.

Simani, S. and M. Benini (2007). Residual Generator Design for the FDI of Linear Multivariable Sampled–Data Dynamic Systems. In: *CDC2007 – 46–th IEEE Conference on Decision and Control* (EUCA IEEE, CSS, Ed.). Vol. 1. IEEE, CSS, EUCA. IEEE, CSS, EUCA. New Orleans, LA, U.S.A.. pp. 2602–2607.

Simani, S., C. Fantuzzi and R. J. Patton (2002). *Model-based fault diagnosis in dynamic systems using identification techniques*. Advances in Industrial Control. first ed.. Springer–Verlag. London, UK. ISBN: 1852336854.

Stevens, B. L. and F. L. Lewis (2003). *Aircraft Control and Simulation*. 2^{nd} ed.. John Wiley and Son.

Titterton, D.H. and J.L. Weston (2005). *Strapdown Inertial Navigation Technology*. IEE Radar, Sonar, Navigation and Avionics. 2nd ed.. IEE. London, UK.

Van Huffel, S. and Lemmerling, P., Eds. (2002). *Total Least Squares and Errors–in–Variables Modeling: Analysis, Algorithms and Applications*. 1st ed.. Springer–Verlag. ISBN: 1402004761.

Watanabe, K. and D. M. Himmelblau (1982). Instrument fault detection in systems with uncertainties. *Int. J. System Sci.* **13**(2), 137–158.

Willsky, A. S. (1976). A survey of design methods for failure detection in dynamic systems. *Automatica* **12**(6), 601–611.

Wünnenberg, J. (1990). Observer–based fault detection in dynamic systems. PhD thesis. University of Duisburg. Duisburg, Germany.

Wünnenberg, J. and P. M. Frank (1987). Sensor fault detection via robust observer. In: *System Fault Diagnosis, Reliability, and Related Knowledge-Based Approaches*. S. Tzafestas et al ed.. Vol. 1. pp. 147–160.

Wünnenberg, J. and P. M Frank (1990). Robust observer–based detection for linear and non–linear systems with application to robot. In: *Proc. of IMACS Annals on Computing & Applied Mathematics MIM–S^2:90*. Brussels.

Zad, Shahin Hashtrudi and Mohammad-Ali Massoumnia (1999). Generic solvability of the failure detection and identification problem. *Automatica* **35**(5), 887–893.

Zhang, Q., F. Campillo, Cérou F. and F. Legland (2005). Nonlinear system fault detection and isolation based on boostrap particle filters. In: *Proc. of 44th IEEE CDC-ECC*. Seville, Spain. pp. 3821–3826.