Research Article

Design and Analysis of Robust Fault Diagnosis Schemes for a Simulated Aircraft Model

M. Benini,[1] M. Bonfè,[1] P. Castaldi,[2] W. Geri,[2] and S. Simani[1]

[1] *Department of Engineering, University of Ferrara, Via Saragat 1, 44100 Ferrara, Italy*
[2] *Aerospace Engineering Faculty, University of Bologna, Via Fontanelle 40, 40136 Forlì, Italy*

Correspondence should be addressed to S. Simani, ssimani@ing.unife.it

Several procedures for sensor fault detection and isolation (FDI) applied to a simulated model of a commercial aircraft are presented. The main contributions of the paper 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). The FDI strategies are applied to the aircraft model, 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 turbulence, measurement, and model errors. Comparisons with other disturbance-decoupling methods for FDI based on neural networks (NNs) and unknown input kalman filter (UIKF) are finally reported.

1. INTRODUCTION

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 [1, 2]. 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. There are, broadly speaking, two main approaches for addressing the FDI problem, namely, hardware-based and model-based techniques [3, 4]. 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 [5], state estimation [6], unknown input observer (UIO), Kalman filters (KFs) [3], and parameter identification [6].Intelligent techniques [7] can be also exploited. Furthermore, the Massoumnia's

geometric method [8] was successfully extended to nonlinear systems [9, 10]. A crucial issue with any FDI scheme is its robustness properties and 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.

The first part of this work deals with the residual generator design for the FDI of input-output sensors of a general aviation aircraft subject to turbulence, wind gust disturbances, and measurement noises. The developed PM scheme belongs to the parity space approach [5] 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. These dynamic filters, organised into bank structures, are able to achieve fault isolation properties. An appropriate choice of their parameters allows to maximise robustness with respect to both measurement noise and modelling errors, while optimising fault sensitivity characteristics. The development of NLGA methodology is based on the works by De Persis and Isidori [10]. 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. This technique was applied for the first time to a vertical takeoff and landing (VTOL) aircraft with reference to a reduced-order model [11]. The NLGA residual generators have been designed in order to be analytically decoupled from the vertical and lateral components of the wind (gusts and turbulence). Moreover, a new full analytical developed mixed $\mathcal{H}_-/\mathcal{H}_\infty$ optimisation is proposed in order to design the NLGA residual generators so that a good tradeoff between the fault sensitivity and the robustness with respect to measurements and model errors is achieved. The designed residual generators have been tested on a PIPER PA-30 aircraft flight simulator that was implemented in Matlab-Simulink environments. With respect to the related works by the same authors [12, 13], the main contribution of this paper regards the enhancement in the designs of the proposed FDI schemes. Moreover, 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 (NNs) 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.

2. AIRCRAFT MODEL OVERVIEW

This section recalls briefly the description of the monitored aircraft whose main parameters and variables are reported in Table 1.

The considered aircraft simulation model consists of a PIPER PA-30, based on the classical nonlinear 6 degrees of freedom (DoF) rigid body formulation [14] whose motion occurs as a consequence of applied forces and moments (aerodynamic, propulsive, and gravitational). A set of local approximations for these forces has been computed and scheduled depending on the values assumed by true airspeed (TAS), curvature radius, flight path angle, altitude, and flap deflection. In this way, it is possible to obtain a mathematical model for each flight condition. This model is 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, 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) but by means of cubic splines approximating the nonlinear experimental curves. The nonlinear 6 DoF model has been completed by means of the PIPER PA-30 propulsion system consisting of two 4-pistons aspirated engines, with the throttle valve aperture δ_{th} as input and the overall thrust intensity as output. The overall simulation model, used to perform all the following tests, consists of the aircraft 6 DoF flight

TABLE 1: Nomenclature.

α	Angle of attack
β	Angle of sideslip
p_ω	Roll rate
q_ω	Pitch rate
r_ω	Yaw rate
ϕ	Bank angle
θ	Elevation angle
ψ	Heading angle
n_e	Engine shaft angular rate
$\begin{bmatrix} I_x & 0 & -I_{xz} \\ 0 & I_y & 0 \\ -I_{xz} & 0 & I_z \end{bmatrix}$	Inertia moment matrix
V	True airspeed (TAS)
δ_e	Elevator deflection angle
δ_a	Aileron deflection angle
δ_r	Rudder deflection angle
δ_{th}	Throttle aperture percentage
H	Altitude
γ	Flight path angle
m	Airplane mass
$\omega_u, \omega_v, \omega_w$	Wind gust components

dynamics and the engine model completed with the model of input-output sensors, the servo actuators, the atmosphere turbulence Dryden description, the wind gust disturbances, and a classical autopilot. Moreover, 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.).

The linear model used by the proposed PM FDI approach described in Section 3 embeds the linearisation both of the 6 DoF model and of the propulsion system as follows:

$$\dot{x}(t) = Ax(t) + Bc(t) + Ed(t) \tag{1}$$

with

$$\begin{aligned} x(t) &= [\Delta V(t)\Delta\alpha(t)\Delta\beta(t)\Delta p_\omega(t)\Delta q_\omega(t)\Delta r_\omega(t) \\ &\quad \cdots \Delta\phi(t)\Delta\theta(t)\Delta\psi(t)\Delta n_e(t)]^T, \\ c(t) &= [\Delta\delta_e(t)\Delta\delta_a(t)\Delta\delta_r(t)\Delta\delta_{th}(t)]^T, \\ d(t) &= [w_u(t)w_v(t)w_w(t)]^T, \end{aligned} \tag{2}$$

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 (1) is of the type $y(t) = Cx(t)$, where the rows of C correspond to rows of the identity matrix, depending on the measured variables.

On the other hand, regarding the NLGA FDI scheme described in Section 4, it requires a nonlinear input affine

system [10], but the adopted simulation model of the aircraft does not fulfil this requirement. For this reason, the following simplified aircraft model is used:

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

where $C_{(\cdot)}$ are the aerodynamic coefficients; $t_{(\cdot)}$ are the engine parameters; and w_v, w_l are the vertical and lateral wind disturbance components. In particular, the model of (3) has been obtained on the basis of some assumptions.In particular, 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. Finally, the rudder effect in the equation describing the β dynamics has been neglected. It is worth noting that Section 5 has shown that the designs and the simulations of the NLGA residual generators are robust with respect to the last approximation. In fact, the model of the β dynamics will never be used.

3. PM RESIDUAL GENERATORS

Let us consider the input-output representation of a continuous-time, time-invariant linear dynamic system affected by faults and disturbances in the form

$$P(s)y(t) = Q_c(s)c(t) + Q_d(s)d(t) + Q_f(s)f(t), \tag{4}$$

where $y(t) \in \mathfrak{R}^m$ is the output vector, $c(t) \in \mathfrak{R}^{l_c}$ is the input vector, $d(t) \in \mathfrak{R}^{l_d}$ is the disturbance vector, and $f(t) \in \mathfrak{R}^{l_f}$ is the fault vector; $P(s)$, $Q_c(s)$, $Q_d(s)$, and $Q_f(s)$ are known polynomial matrices of proper dimensions.

Models of type (4) can be frequently found in practice by applying well-known physical laws to describe the input-output dynamical links of various systems. Algorithms to transform multivariable state-space models to equivalent multiple-input-multiple-output (MIMO) polynomial representations and vice versa are available [15]. Suitable software routines for multivariable system transformations have been implemented by the authors in the Matlab environment. In fact, the Matlab software for state-space and transfer function conversions is not able to manage directly MIMO models, since they are considered as concatenations of single-input-single-output (SISO) systems.

An important aspect of the residual generator design concerns the decoupling properties of the disturbance $d(t)$. The decoupling can be obtained premultiplying all the terms of (4) by the matrix $L(s) \in \mathcal{N}_l(Q_d(s))$, that is, the left null-space of the matrix $Q_d(s)$:

$$L(s)P(s)y(t) - L(s)Q_c(s)c(t) = L(s)Q_f(s)f(t). \tag{5}$$

Hence, the residual generator for the system of (4) is represented by

$$
\begin{aligned}
R(s)r(t) &= L(s)P(s)y(t) - L(s)Q_c(s)c(t) \\
&= L(s)Q_f(s)f(t),
\end{aligned}
\tag{6}
$$

where it is assumed that $r(t) \in \mathfrak{R}$ and $L(s)$ is a polynomial row vector. The polynomial $R(s)$ can be arbitrarily selected among the polynomials with degree greater than or equal to n_r^*, where n_r^* is the maximum row-degree of the pair $\{L(s)P(s), L(s)Q_c(s)\}$. Moreover, if all the roots of $R(s)$ lie in the open left-half s-plane, it assures the stability of the filter of (6). Without loss of generality, it is assumed that $R(0) = 1$.

Remark 1. If the matrix $Q_d(s)$ is of full-column rank (i.e., rank $Q_d(s) = l_d$), $\mathcal{N}_l(Q_d(s))$ has dimension $m - l_d$. Therefore, a polynomial matrix $B(s)$, whose rows represents a minimal polynomial basis of $\mathcal{N}_l(Q_d(s))$, has $m - l_d$ rows and m columns.

This work is focused on the problem of detecting and isolating additive faults acting on the input and output sensors of the monitored system. If the input-output measurements

are modelled by the relations of (7):

$$c^*(t) = c(t) + f_c(t),$$
$$y^*(t) = y(t) + f_o(t),$$
(7)

the system of (4) becomes

$$P(s)(y^*(t) - f_o(t)) = Q_c(s)(c^*(t) - f_c(t)) + Q_d(s)d(t),$$
(8)

under the assumptions that $Q_f(s)f(t) = [-Q_c(s), P(s)] \cdot [f_c^T(t), f_o^T(t)]^T$. Thus, the residual generator of (6) is written as

$$R(s)r(t) = L(s)P(s)y^*(t) - L(s)Q_c(s)c^*(t)$$
$$= L(s)P(s)f_o(t) - L(s)Q_c(s)f_c(t).$$
(9)

Remark 2. The residual generator described by (7) and (9) can be seen as an errors-in-variables (EIV) model [16] with respect the input and output variables, 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 (9).

Remark 3. The diagnostic capabilities of the residual generator of (6) strongly depend on the choice of the terms $L(s)$ and $R(s)$. This paper proposes a method for the design of these polynomials, under the assumption that $f(t)$ is a scalar and, consequently, $Q_f(s)$ is a vector. The rationale of this assumption is commented in Section 3.2 where the fault isolation method is proposed.

In the following, the freedom design in the selection of the rows of the polynomial matrix $L(s)$ is investigated when $q = m - l_d \geq 2$. These degrees of freedom are 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 $G_f(s) = L(s)Q_f(s)/R(s)$.

If $b_i(s)$ $(i = 1, \ldots, q)$ are the row vectors of the basis $B(s)$, $L(s)$ can be expressed as linear combination of these vectors:

$$L(s) = \sum_{i=1}^{q} k_i b_i(s),$$
(10)

where k_i are real constants maximising:

$$\lim_{s \to 0} \frac{1}{R(s)} \left[\sum_{i=1}^{q} k_i b_i(s) \right] Q_f(s) = \left[\sum_{i=1}^{q} k_i b_i(0) \right] Q_f(0)$$
(11)

with the constraint

$$\sum_{i=1}^{q} k_i^2 = 1.$$
(12)

Under these assumptions, when the fault $f(t)$ is a step-function of magnitude F, the steady-state residual value is

$$\lim_{t \to \infty} r(t) = \lim_{s \to 0} s \frac{L(s)Q_f(s)}{R(s)} \frac{F}{s} = \left[\sum_{i=1}^{q} k_i b_i(0) \right] Q_f(0)F.$$
(13)

If the following real vectors are defined as

$$k = \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_q \end{bmatrix}, \qquad a = B(0)Q_f(0) = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_q \end{bmatrix},$$
(14)

the problem of the maximisation of the residual fault sensitivity can be recast as follows.

Proposition 1. *Given the vector a, the vector k that maximises the steady-state fault sensitivity, that is, the function W given by the expression*

$$W = a^T k = \sum_{i=1}^{q} a_i k_i,$$
(15)

under the constraint of (12)*, can be found by solving*

$$\tilde{k} = \arg \max W(k).$$
(16)

The solution to the problem described by Proposition 1 can be derived as follows. The constraint of (12) describes a hypersphere, whilst the expression of the function of (15) is a hyperplane. The unknown coefficients k 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 W. As shown below, the solution of the problem described by Proposition 1 exists and is unique.

Proof. From (12), k_1 is expressed as a function of k_2, $k_3, \ldots, k_q$, and it is substituted into (15):

$$W = a_1 \sqrt{1 - k_2^2 - k_3^2 - \cdots - k_q^2} + a_2 k_2 + \cdots + a_q k_q.$$
(17)

By computing $\nabla W = 0$, that is,

$$\frac{\partial W}{\partial k_2} = \frac{1}{2} a_1 \frac{-2k_2}{\sqrt{1 - k_2^2 - k_3^2 - \cdots - k_q^2}} + a_2 = 0,$$

$$\frac{\partial W}{\partial k_3} = \frac{1}{2} a_1 \frac{-2k_3}{\sqrt{1 - k_2^2 - k_3^2 - \cdots - k_q^2}} + a_3 = 0,$$
(18)

$$\vdots$$

$$\frac{\partial W}{\partial k_q} = \frac{1}{2} a_1 \frac{-2k_q}{\sqrt{1 - k_2^2 - k_3^2 - \cdots - k_q^2}} + a_q = 0,$$

and squaring the expression, after algebraic manipulation:

$$a_2^2 = (a_2^2 + a_1^2)k_2^2 + a_2^2 k_3^2 + \cdots + a_2^2 k_q^2,$$

$$a_3^2 = a_3^2 k_2^2 + (a_3^2 + a_1^2)k_3^2 + \cdots + a_3^2 k_q^2,$$
(19)

$$\vdots$$

$$a_q^2 = a_q^2 k_2^2 + a_q^2 k_3^2 + \cdots + (a_q^2 + a_1^2)k_q^2,$$

an expression in the form of $Ax = b$ is obtained, where

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

$$
x = \begin{bmatrix} k_2^2 \\ k_3^2 \\ \vdots \\ k_q^2 \end{bmatrix}, \qquad b = \begin{bmatrix} a_2^2 \\ a_3^2 \\ \vdots \\ a_q^2 \end{bmatrix}.
\tag{20}
$$

The unknown vector $\tilde{x}$, under the constraint of (12), can be expressed as follows:

$$
\tilde{x} = \begin{bmatrix} 1 - \sum_{i=1}^{q-1} (A^{-1}b)_i \\ A^{-1}b \end{bmatrix},
\tag{21}
$$

where $(A^{-1}b)_i$ is the ith element of the vector $A^{-1}b$. The vector $\tilde{x}$ represents the squares of the solution of the problem of Proposition 1.

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 Proposition 1 can be reformulated as

$$
\tilde{k} = \arg \max_{k \in \Omega} W(k).
\tag{22}
$$

$\square$

Remark 4. 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 to reapply the same procedure.

Remark 5. The same solution can be found by maximising the function $|W|$. Due to the symmetry properties, the maximisation of $|W|$ admits two solutions corresponding to the maximum and the minimum of the function W. Moreover, the choice of the quadratic constraint of (12) guarantees the unicity of the solution to the problem of Proposition 1.

Remark 6. The problem described by Proposition 1 could have been solved also in a numerical way, that is, by searching k that maximises W on the surface of the q-dimensional hypersphere. However, the computational cost of this numerical solution can be a drawback when q is big.

3.1. PM residual design

Section 3 has shown how to maximise the steady-state gain of the continuous-time transfer function $G_f(s) = L(s)Q_f(s)/R(s)$ trough a suitable choice of the real vector k

(i.e., $k = \tilde{k}$). The design of the filter of (6) has been completed here by introducing a method for assigning both the zeros and the poles of the continuous-time transfer function $G_f(s)$. The zeros and poles location influences the transient characteristics (maximum overshoot, delay time, rise time, settling time, etc.) of the filter of (6). In many applications, these characteristics must be kept within tolerable or prescribed limits in order to guarantee good performances of the filter in terms, for example, of fault detection times and false-alarm probabilities.

Remark 7. When $k = \tilde{k}$, the polynomial $L(s)Q_f(s) = k^T B(s)Q_f(s)$ is fixed and no freedom degree is left to arbitrarily assign the zeros. In order to solve this problem, a polynomial vector $k(s)$ can be considered. Under this assumption, $L(s)$ still belongs to the subspace $\mathcal{N}_l(Q_d(s))$, where the terms k_i are polynomial coefficients.

The previous consideration leads to introduce the polynomial $E(s) = k^T(s)B(s)Q_f(s)$, where $k(s)$ is a q-dimensional polynomial vector whose ith element has the form

$$
k_i(s) = \sum_{j=0}^{n_k} k_i^j s^j.
\tag{23}
$$

The degree n_k and the $q \times n_k$ coefficients k_i^j are freedom design ($j \neq 0$) that are exploited for obtaining the desired roots of the polynomial $E(s)$. However, in order to maximise the steady-state gain, as shown in Section 3, the following condition must hold:

$$
k(0) = \tilde{k} = \begin{bmatrix} \tilde{k}_1 \\ \tilde{k}_2 \\ \vdots \\ \tilde{k}_q \end{bmatrix} \iff k_i^0 = \tilde{k}_i, \quad i = 1, \ldots, q.
\tag{24}
$$

Definition 1. $H(s)$ is the reference polynomial whose roots are the zeros to be assigned:

$$
H(s) = \sum_{j=0}^{n_h} h^j s^j.
\tag{25}
$$

Since the constraint of (24) must hold, $H(0) = \tilde{k}^T B(0) \cdot Q_f(0)$. Obviously, this assumption does not provide any restriction on the roots assignable. Under the previous considerations, the zero assignment and pole placement problem is formulated as follows.

Proposition 2. *The degree n_k and the coefficients k_i^j have to be determined under the constraint of (24) in order to obtain $E(s) = H(s)$.*

Proof. In Section 3, the polynomial vector $a(s) = B(s)Q_f(s)$ was defined. Its ith element is a known polynomial of a

certain degree, n_{a_i}. If n_a is defined as follows:

$$n_a = \max_{i=1,\dots,q} n_{a_i}, \qquad (26)$$

the ith element of $a(s)$ can be always written as a polynomial of degree n_a:

$$a_i(s) = \sum_{j=0}^{n_a} a_i^j s^j \qquad (27)$$

by imposing that $a_i^j = 0$ when $j > n_{a_i}$.

As $E(s) = k^T(s)a(s)$, by multiplying (23) and (27), it results

$$E(s) = \sum_{i=1}^{q} \sum_{j=0}^{n_k+n_a} \left(\sum_{\alpha+\beta=j} k_i^\alpha a_i^\beta \right) s^j = \sum_{j=0}^{n_k+n_a} e^j s^j, \qquad (28)$$

where

$$e^j = \sum_{i=1}^{q} \sum_{\alpha+\beta=j} k_i^\alpha a_i^\beta. \qquad (29)$$

Equations (28) and (29) assume that $k_i^\alpha = 0$ when $\alpha > n_k$ and $a_i^\beta = 0$ when $\beta > n_a$. Note that the coefficients $e^1, \dots, e^{n_k+n_a}$ depend on the freedom design $k_i^1, \dots, k_i^{n_k}$. On the other hand, e^0 is fixed as the coefficients k_i^0 are assigned by (24).

Let us suppose that $n_h \le n_k + n_a$. By imposing $E(s) = H(s)$, from (29) and (25), the following expressions are computed:

$$\sum_{i=1}^{q} \sum_{\alpha+\beta=j\,(\alpha\neq 0)} k_i^\alpha a_i^\beta = h^j - \sum_{i=1}^{q} k_i^0 a_i^j, \quad j = 1,\dots,n_k+n_a. \qquad (30)$$

Equations (24) and (30) represent a linear system with n_k+n_a equations and $q \times n_k$ unknowns, which can be expressed in the classical form $Ax = b$, where

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

$$x = \begin{bmatrix} k_1^1 \\ \vdots \\ k_q^1 \\ k_1^2 \\ \vdots \\ k_q^2 \\ \vdots \\ \vdots \\ k_1^{n_k} \\ \vdots \\ k_q^{n_k} \end{bmatrix}, \qquad b = \begin{bmatrix} h^1 - \sum_{i=1}^{q} k_i^0 a_i^1 \\ \vdots \\ h^{n_a} - \sum_{i=1}^{q} k_i^0 a_i^{n_a} \\ h^{n_a+1} \\ \vdots \\ h^{n_a+n_k} \end{bmatrix}. \qquad (31)$$

The degree n_k of the polynomials $k_i(s)$ has to be chosen in order to obtain a solvable system (i.e., rank A = rank $[A\ b]$). □

In order to understand the proposed solution, the following points should be considered.

(i) The choice of n_k must guarantee that the relations $n_h \le n_k + n_a$ are satisfied.

(ii) When $q \ge 2$, the difference between the number of unknown terms and the number of equations, that is, $(q - 1) \times n_k - n_a$, is greater than zero if n_k is selected sufficiently high.

(iii) 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 pseudoinverse of A, that is, $A^+ b$ can be computed.

Remark 8. 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 that

$$\frac{G_f(s)}{G_f(0)} = \frac{F(0)}{F(s)}, \qquad (32)$$

where $F(s)$ is an arbitrary polynomial. These cases do not require a $k(s)$ such that $E(s) = G_f(0)$, but it is enough considering $k = \tilde{k}$ and imposing

$$R(s) = \frac{E_0(s)F(s)}{G_f(0)F(0)}, \qquad (33)$$

where $E_0(s) = \tilde{k}B(s)Q_f(s)$. However, there is a restriction on the choice of $F(s)$. In fact, due to the realisability condition, $\deg\{F(s)\} > n_r^* - \deg\{E_0(s)\}$. Moreover, the method cannot be applied if $E_0(s)$ admits one or more roots with positive real part, as the residual generator would become unstable. These cases require an approximate solution.

Remark 9. This section 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, for example, in a work by the same authors [17].

3.2. PM fault isolation

This section addresses the design problem of residual generator banks for the isolation of faults affecting the input and the output sensors. This design is performed by using the disturbance-decoupling method suggested in Section 3.

To univocally isolate a fault concerning one of the *output sensors*, under the hypotheses that the input sensors and the remaining output sensors are fault-free, a bank of residual generator filters is used. The number of these generators is equal to the number m of the system outputs, and the ith device ($i = 1, \ldots, m$) is driven by all but the ith output and all the inputs of the system. In this case, a fault on the ith output sensor affects all but the ith residual generator.

In presence of a fault on the jth output sensor, the measured output $y^*(t)$ can be expressed as follows:

$$y^*(t) = y(t) + f_{o_j}(t), \qquad (34)$$

with

$$f_{o_j}(t) = \begin{bmatrix} 0 & \cdots & 0 & h_{o_j}(t) & 0 & \cdots & 0 \end{bmatrix}^T, \qquad (35)$$

where $h_{o_j}(t)$ represents the jth output fault function.

In these conditions, the system of (4) becomes

$$P(s)y^*(t) - p_j(s)h_{o_j}(t) = Q_c(s)c(t) + Q_d(s)d(t), \qquad (36)$$

where $p_j(s)$ is the jth column of the matrix $P(s)$.

Let us indicate $L_{o_i}(s)$ a polynomial row vector belonging to the basis of the left null space of the matrix $[Q_d(s) \mid p_i(s)]$. The expression of the ith filter when a fault is acting on the jth output sensor is obtained by multiplying (36) by $L_{o_i}(s)$:

$$R_{o_i}(s)r_{o_i}(t) = L_{o_i}(s)P^i(s)y^{*i}(t) - L_{o_i}(s)Q_c(s)c(t)$$
$$= \begin{cases} L_{o_i}(s)p_j(s)h_{o_j}(t) & \text{for } j \neq i, \\ 0 & \text{for } j = i, \end{cases} \qquad (37)$$

where $P^i(s)$ is the matrix obtained by deleting from $P(s)$ the ith column, and $y^{*i}(t)$ represents the $(m-1)$-dimensional vector obtained by deleting from $y^*(t)$ its ith component.

From the comparison between (37) and (6) with $f(t) \in \mathfrak{R}$ if $q' = m - l_d - 1 \geq 2$, the methods shown in Sections 3 and 3.1 can be exploited for the design of the ith filter. In particular, the parameters of this filter can be properly chosen in order to optimise its performances when a fault is acting on the jth output sensor.

In more detail, as shown in Section 3, $L_{o_i}(s)$ is chosen to maximise the steady-state gain in the presence of the fault $f_{o_j}(t)$. Moreover, as shown in Section 3.1, $R_{o_i}(s)$ is chosen to obtain a fixed behaviour of the transfer function due to the fault $f_{o_j}(t)$.

It is worth noting that the similar design technique can be used for *input sensor* fault isolation.

The problem requirements determine the selection of the specific fault with respect to which the design depends. Most often in practice, it is important to obtain *good* performance with respect to all possible faults rather than *optimal* behaviour with respect to one specific fault. In this situation, a different design of the filter behaviour for each fault situation is needed.

4. NLGA RESIDUAL GENERATORS

The considered NLGA to the FDI problem is suggested in [18] and formally developed in [10]. 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.

More precisely, the approach considers a nonlinear system model in the form

$$\begin{aligned} \dot{x} &= n(x) + g(x)c + \ell(x)f + p(x)d, \\ y &= h(x) \end{aligned} \qquad (38)$$

in which the state vector $x \in \mathcal{X}$ (an open subset of $\mathfrak{R}^{\ell_n}$), $c(t) \in \mathfrak{R}^{\ell_c}$ is the control input vector, $f(t) \in \mathfrak{R}$ is the fault, $d(t) \in \mathfrak{R}^{\ell_d}$ the disturbance vector (embedding also the faults which have to be decoupled), and $y \in \mathfrak{R}^{\ell_m}$ the output vector; whilst $n(x)$, $\ell(x)$, 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 devised as it follows: first, determine the largest observability codistribution contained in $P^\perp$, denoted with Ω^* [10].

If $\ell(x) \notin \Omega^*$, the design procedure can continue, otherwise, the fault is not detectable; whenever the previous condition 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

$$\Psi(y) = \begin{pmatrix} \overline{y}_1 \\ \overline{y}_2 \end{pmatrix} = \begin{pmatrix} \Psi_1(y) \\ H_2 y \end{pmatrix}, \qquad \Phi(x) = \begin{pmatrix} \overline{x}_1 \\ \overline{x}_2 \\ \overline{x}_3 \end{pmatrix} = \begin{pmatrix} \Phi_1(x) \\ H_2 h(x) \\ \Phi_3(x) \end{pmatrix} \qquad (39)$$

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 of (38) is described by the relations in the form

$$
\begin{aligned}
\dot{\bar{x}}_1 &= n_1(\bar{x}_1, \bar{x}_2) + g_1(\bar{x}_1, \bar{x}_2)c + \ell_1(\bar{x}_1, \bar{x}_2, \bar{x}_3)f, \\
\dot{\bar{x}}_2 &= n_2(\bar{x}_1, \bar{x}_2, \bar{x}_3) + g_2(\bar{x}_1, \bar{x}_2, \bar{x}_3)c \\
&\quad + \ell_2(\bar{x}_1, \bar{x}_2, \bar{x}_3)f + p_2(\bar{x}_1, \bar{x}_2, \bar{x}_3)d, \\
\dot{\bar{x}}_3 &= n_3(\bar{x}_1, \bar{x}_2, \bar{x}_3) + g_3(\bar{x}_1, \bar{x}_2, \bar{x}_3)c \\
&\quad + \ell_3(\bar{x}_1, \bar{x}_2, \bar{x}_3)f + p_3(\bar{x}_1, \bar{x}_2, \bar{x}_3)d, \\
\bar{y}_1 &= h(\bar{x}_1), \\
\bar{y}_2 &= \bar{x}_2
\end{aligned}
\tag{40}
$$

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:

$$
\begin{aligned}
\dot{\bar{x}}_1 &= n_1(\bar{x}_1, \bar{y}_2) + g_1(\bar{x}_1, \bar{y}_2)c + \ell_1(\bar{x}_1, \bar{y}_2, \bar{x}_3)f, \\
\bar{y}_1 &= h(\bar{x}_1),
\end{aligned}
\tag{41}
$$

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.

As already described in Section 2, the proposed NLGA FDI scheme is designed on the basis of a model structure of the input affine type as described in [10]. For this reason, the aircraft simulation model has been simplified and the nonlinear model of (3) has been considered for the NLGA design.

Under these assumptions, by means of computations detailed in [19], the residual generators for detecting the faults affecting the aircraft input sensors are obtained. In particular, the residual generator for the elevator $r_{\delta_e}(t)$, with $k_{\delta_e} > 0$, is described by the relation

$$
\begin{aligned}
\dot{\xi}_1 &= \frac{V^2}{m}\left[-(C_{D0} + C_{D\alpha}\alpha + C_{D\alpha_2}\alpha^2)\cos\alpha \right] \\
&\quad + \frac{V^2}{m}(C_{L0} + C_{L\alpha}\alpha)\sin\alpha - g\sin\theta \\
&\quad + Vq_\omega\sin\alpha - \frac{(C_{m0} + C_{m\alpha}\alpha + C_{mq}q_\omega)}{mt_d}V^2 \\
&\quad - \frac{(I_z - I_x)}{mt_d}p_\omega r_\omega - \frac{C_{\delta_e}}{mt_d}V^2\delta_e \\
&\quad + k_{\delta_e}\left[\left(V\cos\alpha - \frac{I_y}{mt_d}q_\omega\right) - \xi_1\right], \\
r_{\delta_e} &= \left(V\cos\alpha - \frac{I_y}{mt_d}q_\omega\right) - \xi_1.
\end{aligned}
\tag{42}
$$

The aileron residual generator $r_{\delta_a}(t)$, with $k_{\delta_a} > 0$, has the form

$$
\begin{aligned}
\dot{\xi}_2 &= \frac{(C_{l\beta}\beta + C_{lp}p_\omega)}{I_x}V^2 + \frac{(I_y - I_z)}{I_x}q_\omega r_\omega \\
&\quad + \frac{C_{\delta_a}}{I_x}V^2\delta_a + k_{\delta_a}(p_\omega - \xi_2), \\
r_{\delta_a} &= p_\omega - \xi_2.
\end{aligned}
\tag{43}
$$

The rudder residual generator $r_{\delta_r}(t)$, with $k_{\delta_r} > 0$, is written in the form

$$
\begin{aligned}
\dot{\xi}_3 &= \frac{(C_{n\beta}\beta + C_{nr}r_\omega)}{I_z}V^2 + \frac{(I_x - I_y)}{I_z}p_\omega q_\omega \\
&\quad + \frac{C_{\delta_r}}{I_z}V^2\delta_r + k_{\delta_r}(r_\omega - \xi_3), \\
r_{\delta_r} &= r_\omega - \xi_3.
\end{aligned}
\tag{44}
$$

The throttle residual generator $r_{\delta_{th}}(t)$, with $k_{\delta_{th}} > 0$, has the form

$$
\begin{aligned}
\dot{\xi}_4 &= t_n n_e^3 + \frac{t_f}{n_e}(t_0 + t_1 n_e)\delta_{th} + k_{\delta_{th}}(n_e - \xi_4), \\
r_{\delta_{th}} &= n_e - \xi_4.
\end{aligned}
\tag{45}
$$

Remark 10. It is worth observing that 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. This feature can be obtained by defining a different $p(x)$ for each residual generator design [19]. 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 dynamics (42)–(45).

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

4.1. NLGA robustness improvement

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

(1) the structural decoupling of critical disturbances (wind gust and turbulence) and critical modelling errors can be obtained as described in Section 4;

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

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 (41). In particular, the vectors $\bar{x}_1$ and $\bar{y}_1$ are decomposed as follows:

$$
\bar{x}_1 = \begin{bmatrix} \bar{x}_{11} \\ \bar{x}_{1c} \end{bmatrix}, \qquad \bar{y}_1 = \begin{bmatrix} \bar{y}_{11} \\ \bar{y}_{1c} \end{bmatrix},
\tag{46}
$$

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

$$n_1(\cdot) = \begin{bmatrix} n_{11}(\cdot) \\ n_{1c}(\cdot) \end{bmatrix}, \quad g_1(\cdot) = \begin{bmatrix} g_{11}(\cdot) \\ g_{1c}(\cdot) \end{bmatrix}, \quad \ell_1(\cdot) = \begin{bmatrix} \ell_{11}(\cdot) \\ \ell_{1c}(\cdot) \end{bmatrix} \tag{47}$$

Let us consider the following conditions:

$$\overline{y}_{11} = h_{11}(\overline{x}_{11}) \qquad \overline{y}_{1c} = h_{1c}(\overline{x}_{1c}) \qquad \ell_{11}(\cdot) \neq 0 , \tag{48}$$

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 (48) are satisfied, the decomposition (46)-(47) can always be applied to obtain the following $\overline{x}_{11}$-subsystem:

$$\begin{aligned} \dot{\overline{x}}_{11} &= n_{11}(\overline{x}_{11}, \overline{y}_{1c}, \overline{y}_2) + g_{11}(\overline{x}_{11}, \overline{y}_{1c}, \overline{y}_2)c \\ &\quad + \ell_{11}(\overline{x}_{11}, \overline{y}_{1c}, \overline{y}_2, \overline{x}_3)f, \\ \overline{y}_{11} &= h_{11}(\overline{x}_{11}). \end{aligned} \tag{49}$$

As can be seen in [19], the conditions (48) are satisfied for the considered aircraft application, hence, from now on, the scalar $\overline{x}_{11}$-subsystem (49) is referred to in place of the $\overline{x}_1$-subsystem (41).

It can be noted that the tuning of the residual generator gains, in the framework of the $\overline{x}_{11}$-subsystem (49), cannot be properly carried out. In fact, the critical disturbances are structurally decoupled but the noncritical 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 noncritical disturbances in the $\overline{x}_{11}$-subsystem as follows:

$$\begin{aligned} \dot{\overline{x}}_{11} &= n_{11}(\overline{x}_{11}, \overline{y}_{1c}, \overline{y}_2) + g_{11}(\overline{x}_{11}, \overline{y}_{1c}, \overline{y}_2)c \\ &\quad + \ell_{11}(\overline{x}_{11}, \overline{y}_{1c}, \overline{y}_2, \overline{x}_3)f + e(\overline{x}_{11}, \overline{y}_{1c}, \overline{y}_2, \overline{x}_3)\zeta, \\ \overline{y}_{11} &= \overline{x}_{11} + \nu, \end{aligned} \tag{50}$$

where, to simplify the treatment without loss of generality (accordingly to the considered aircraft application), the state variable $\overline{x}_{11}$ is supposed to be directly measured. Moreover, the variable $\nu \in \Re$ is the measurement noise on $\overline{x}_{11}$. Finally, the variable $\zeta \in \Re$ and the related scalar field $e(\cdot)$ represent the noncritical effects which have not been considered in the simplified aircraft model (3) used for the NLGA scheme.

The following system, which is referred to as *filter form*, represents a generic scalar residual generator (based on the subsystem (50)) to which (42)–(45) belong as a particular case

$$\begin{aligned} \dot{\xi}_f &= n_{11}(\overline{y}_{11}, \overline{y}_{1c}, \overline{y}_2) + g_{11}(\overline{y}_{11}, \overline{y}_{1c}, \overline{y}_2)c + k_f(\overline{y}_{11} - \xi_f), \\ r_f &= \overline{y}_{11} - \xi_f, \end{aligned} \tag{51}$$

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

$$\tilde{x}_f = \overline{x}_{11} - \xi_f, \tag{52}$$

which allows to write the following equivalent residual model:

$$\begin{aligned} \dot{\tilde{x}}_f &= n_{11}(\overline{x}_{11}, \overline{y}_{1c}, \overline{y}_2) - n_{11}(\overline{y}_{11}, \overline{y}_{1c}, \overline{y}_2) \\ &\quad + g_{11}(\overline{x}_{11}, \overline{y}_{1c}, \overline{y}_2)c - g_{11}(\overline{y}_{11}, \overline{y}_{1c}, \overline{y}_2)c \\ &\quad + \ell_{11}(\overline{x}_{11}, \overline{y}_{1c}, \overline{y}_2, \overline{x}_3)f + e(\overline{x}_{11}, \overline{y}_{1c}, \overline{y}_2, \overline{x}_3)\zeta \\ &\quad - k_f\tilde{x}_f - k_f\nu, \\ r_f &= \tilde{x}_f + \nu. \end{aligned} \tag{53}$$

In order to apply the effective mixed $\mathcal{H}_-/\mathcal{H}_\infty$ approach [3, 20] to tune k_f, the system (53) has to be linearised in the neighbourhood of a stationary flight condition, as suggested in [2] 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 $\overline{x}_{10}$, $\overline{x}_{30}$, c_0, $\overline{y}_{10}$, and $\overline{y}_{20}$, hence the robustness of the nonlinear residual generator is achieved. The linearisation of (53) is the following:

$$\begin{aligned} \dot{\tilde{x}}_f &= -k_f\tilde{x}_f - k_f\nu + mf + \breve{q}\breve{\zeta}, \\ r_f &= \tilde{x}_f + \nu, \end{aligned} \tag{54}$$

where

$$\begin{aligned} a' &= \left. \frac{\partial n_{11}(\cdot)}{\partial \overline{x}_{11}} \right|_{(\overline{x}_{10}, \overline{y}_{20})}, \qquad b = g_{11}(\cdot)|_{(\overline{x}_{10}, \overline{y}_{20})}, \\ m &= \ell_{11}(\cdot)|_{(\overline{x}_{10}, \overline{y}_{20}, \overline{x}_{30})}, \qquad q = e(\cdot)|_{(\overline{x}_{10}, \overline{y}_{20}, \overline{x}_{30})}, \\ \breve{q}\breve{\zeta} &= q\zeta - a'\nu. \end{aligned} \tag{55}$$

Now, it is important to note that in place of the residual generators in the *filter form* (51), the following *observer form* of the residual generators can be used:

$$\begin{aligned} \dot{\xi}_o &= n_{11}(\xi_o, \overline{y}_{1c}, \overline{y}_2) + g_{11}(\xi_o, \overline{y}_{1c}, \overline{y}_2)c + k_o(\overline{y}_{11} - \xi_o), \\ r_o &= \overline{y}_{11} - \xi_o. \end{aligned} \tag{56}$$

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

$$\tilde{x}_o = \overline{x}_{11} - \xi_o, \tag{57}$$

hence

$$\begin{aligned} \dot{\tilde{x}}_o &= n_{11}(\overline{x}_{11}, \overline{y}_{1c}, \overline{y}_2) - n_{11}(\xi_o, \overline{y}_{1c}, \overline{y}_2) \\ &\quad + g_{11}(\overline{x}_{11}, \overline{y}_{1c}, \overline{y}_2)c - g_{11}(\xi_o, \overline{y}_{1c}, \overline{y}_2)c \\ &\quad + \ell_{11}(\overline{x}_{11}, \overline{y}_{1c}, \overline{y}_2, \overline{x}_3)f + e(\overline{x}_{11}, \overline{y}_{1c}, \overline{y}_2, \overline{x}_3)\zeta \\ &\quad - k_o\tilde{x}_o - k_o\nu, \\ r_o &= \tilde{x}_o + \nu. \end{aligned} \tag{58}$$

The linearisation of (58) is

$$\begin{aligned} \dot{\tilde{x}}_o &= (a' - k_o)\tilde{x}_o - k_o\nu + mf + q\zeta, \\ r_o &= \tilde{x}_o + \nu. \end{aligned} \tag{59}$$

Both the linearised models (54) and (59) of the residual generators in the *filter form* and *observer form*, respectively, can be represented by the following *general form*:

$$\dot{\tilde{x}} = (a - k)\tilde{x} + (E_1 - kE_2)\varepsilon + \mathrm{m}f,$$
$$r = \tilde{x} + E_2\varepsilon \tag{60}$$

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

general form	$\tilde{x}$	ε	r	a	k	e_{11}	E_2
filter form	$\tilde{x}_f$	$\begin{bmatrix} \check{\zeta} \\ \nu \end{bmatrix}$	r_f	0	k_f	$\check{q}$	$\begin{bmatrix} 0 & 1 \end{bmatrix}$
observer form	$\tilde{x}_o$	$\begin{bmatrix} \zeta \\ \nu \end{bmatrix}$	r_o	a	k_o	q	$\begin{bmatrix} 0 & 1 \end{bmatrix}$

(61)

On the basis of (60) and (61), the mixed $\mathcal{H}_-/\mathcal{H}_\infty$ [3, 20] 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.

The following definition will be used throughout the section.

Definition 2. The norms $\mathcal{H}_\infty$ and $\mathcal{H}_-$ of a stable transfer function G are defined as

$$\|G\|_\infty = \sup_{\omega \geq 0} \bar{\sigma}[G(j\omega)], \qquad \|G\|_- = \underline{\sigma}[G(j0)], \tag{62}$$

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

Problem 1 (Mixed $\mathcal{H}_-/\mathcal{H}_\infty$ residual robustness improvement). Given two scalars $\beta > 0$ and $\gamma > 0$, find the set $\mathcal{K}$ defined as:

$$\mathcal{K} = \{k \in \mathfrak{R} : (a - k) < 0, \|G_{r\varepsilon}\|_\infty < \gamma, \|G_{rf}\|_- > \beta\}, \tag{63}$$

where

$$G_{r\varepsilon}(s) = (s - a + k)^{-1}(E_1 - kE_2) + E_2, \tag{64}$$

$$G_{rf}(s) = (s - a + k)^{-1}m. \tag{65}$$

In order to obtain the analytical solution of Problem 1, the following propositions are given.

Proposition 3. *For all $k \in \mathfrak{R}$, $(a - k) < 0$, then*

$$\|G_{r\varepsilon}\|_\infty^2 = \max\left\{1, \frac{(e_{11}^2 + a^2)}{(k - a)^2}\right\}, \tag{66}$$

$$\sup_{\{k \in \mathfrak{R} : (a-k)<0\}} \|G_{r\varepsilon}\|_\infty = +\infty. \tag{67}$$

Proof. From the definition (64)

$$G_{r\varepsilon}(s) = \begin{bmatrix} \dfrac{e_{11}}{s - a + k} & \dfrac{s - a}{s - a + k} \end{bmatrix}, \tag{68}$$

hence it is possible to write

$$\{\bar{\sigma}[G_{r\varepsilon}(j\omega)]\}^2 = \frac{e_{11}^2}{(k - a)^2 + \omega^2} + \frac{a^2 + \omega^2}{(k - a)^2 + \omega^2}$$
$$= \frac{(e_{11}^2 + a^2) + \omega^2}{(k - a)^2 + \omega^2} \tag{69}$$

so that it follows

$$\|G_{r\varepsilon}\|_\infty^2 = \sup_{\xi \geq 0} \frac{(e_{11}^2 + a^2) + \xi}{(k - a)^2 + \xi}. \tag{70}$$

From the last expression, it is straightforward to obtain (66) and (67). □

Proposition 4. *The set*

$$\mathcal{K}_\gamma = \{k \in \mathfrak{R} : (a - k) < 0, \|G_{r\varepsilon}\|_\infty < \gamma, \gamma > 1\} \tag{71}$$

is given by

$$k > \underline{k} \quad \text{with } \underline{k} = a + \frac{\sqrt{e_{11}^2 + a^2}}{\gamma}. \tag{72}$$

Proof. By means of Proposition 3, it is possible to write

$$\frac{(e_{11}^2 + a^2)}{(k - a)^2} < \gamma^2, \tag{73}$$

which holds for

$$k > a + \frac{\sqrt{e_{11}^2 + a^2}}{\gamma}. \tag{74}$$

□

Proposition 5. *If $\gamma > 1$, then $\{\|G_{rf}\|_- : \|G_{r\varepsilon}\|_\infty < \gamma\}$ is given by*

$$0 < \|G_{rf}\|_- < \beta_{\max}(\gamma) \quad \text{where } \beta_{\max}(\gamma) = \frac{m\gamma}{\sqrt{e_{11}^2 + a^2}}. \tag{75}$$

Proof. From the definition (65), 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 4, the maximum feasible value of β fulfilling the constraint $\|G_{r\varepsilon}\|_\infty < \gamma$ is given by

$$\underline{k} = a + \frac{m}{\beta_{\max}(\gamma)}, \tag{76}$$

hence

$$\beta_{\max}(\gamma) = \frac{m}{\underline{k} - a} = \frac{m\gamma}{\sqrt{e_{11}^2 + a^2}}. \tag{77}$$

□

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

$$\mathcal{K} = \left\{ k \in \mathcal{R} : k \in]\underline{k}, \overline{k}[, \; \underline{k} = a + \frac{m}{\beta_{\max}(\gamma)}, \; \overline{k} = a + \frac{m}{\beta} \right\}. \tag{78}$$

Proof. The proof of the theorem is not reported, as it is straightforward from Propositions 3, 4, and 5. □

Remark 11. Let us consider the following performance index to maximise

$$J = \frac{\| G_{rf} \|_-}{\| G_{r\varepsilon} \|_\infty}. \tag{79}$$

From (66), it follows

$$\| G_{r\varepsilon} \|_\infty = \begin{cases} 1, & k > \left(a + \sqrt{e_{11}^2 + a^2} \right) \\ \dfrac{\sqrt{e_{11}^2 + a^2}}{k - a}, & a < k \leq \left(a + \sqrt{e_{11}^2 + a^2} \right) \end{cases} \tag{80}$$

hence

$$J = \begin{cases} \dfrac{m}{k - a}, & k > \left(a + \sqrt{e_{11}^2 + a^2} \right), \\ \dfrac{m}{\sqrt{e_{11}^2 + a^2}}, & a < k \leq \left(a + \sqrt{e_{11}^2 + a^2} \right). \end{cases} \tag{81}$$

From (81), it can be observed that

$$J = \frac{m}{k - a} < \frac{m}{\sqrt{e_{11}^2 + a^2}}, \quad k > \left(a + \sqrt{e_{11}^2 + a^2} \right). \tag{82}$$

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

$$J_{\max} = \frac{m}{\sqrt{e_{11}^2 + a^2}} \quad \forall k \in \mathcal{K}_{\mathcal{J}} = \left\{ k \in \mathcal{R} : a < k \leq \left(a + \sqrt{e_{11}^2 + a^2} \right) \right\}. \tag{83}$$

The method proposed in this paper 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

$$\| G_{rf} \|_- = \frac{m}{k - a} > \beta \geq \frac{m}{\sqrt{e_{11}^2 + a^2}}, \tag{84}$$

hence $k < \left(a + \sqrt{e_{11}^2 + a^2} \right)$.

Finally, from (75) it is always possible to find a β such that

$$\frac{m}{\sqrt{e_{11}^2 + a^2}} \leq \beta \leq \beta_{\max}(\gamma) \quad \forall \gamma > 1. \tag{85}$$

On the basis of Theorem 1, k can be designed by means of the following procedure.

Procedure 1. (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 $k \in]\underline{k}, \overline{k}[$, 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 (51) or to the k_o of (56) if the NLGA residual generator is in the *filter form* or in the *observer form*, respectively.

5. FDI PERFORMANCE ESTIMATION

To show the diagnostic characteristics brought by the application of the proposed FDI schemes to general aviation aircrafts, some numerical results obtained in the Matlab and Simulink environments are reported. The final performances that are achieved with the developed FDI schemes are finally reported. These performances are evaluated by means of extensive simulations applied to the aircraft simulation model. This section presents also some comparisons of the developed PM and NLGA FDI strategies with NN and UIKF FDI schemes.

The designed 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 simulation aircraft model previously described. In particular, a bank of 4 residual generator filters has been used to detect input sensor faults regarding the 4 input control variables. Moreover, in order to obtain the fault isolation properties, each residual generator function of the considered bank is fed by all but one of the 4 control input signals and by the 9 output variables. Obviously, the residual generator bank has been designed to be decoupled from the 3 component wind disturbance vector $d(t) = [w_u(t), w_v(t), w_w(t)]^T$. As to the NLGA residual generator filters, the aircraft synthesis model of (3), adopted for the design, is simplified with respect to the simulation model. Analogously to the PM, the approximations of the NLGA synthesis nonlinear model are related to a particular steady-state flight condition. For this reason, the switching for the NLGA FDI scheme is also required when a generic reference trajectory is followed. Hence, it is important to evaluate the robustness characteristic of a single design of NLGA residual generators when a large set of flight conditions is dealt with.

It is worth noting that the aircraft reference trajectories are typically made up of a sequence of steady-state flight conditions, each one described by the associated input state output set point and the linearised model of (1). As a consequence, all the FDI linear techniques are usually implemented by switching among the residual generators related to the different steady-state flight conditions. The target of this work is to reduce the switching by adopting robust PM residual generators. In particular, the robustness is achieved by using the same residual generators for a large set of flight conditions. The chosen single steady-state flight condition for designing both of the PM and of the NLGA residual

generators is a coordinated turn characterised as follows:

   (i) the true airspeed is 50 m/s;

   (ii) the curvature radius is 1000 m;

   (iii) the flight-path angle is $0°$;

   (iv) the altitude is 330 m;

   (v) the flap deflection is $0°$.

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.

Regarding the PM, the detection properties of the filters in terms of fault sensitivity and disturbance rejection can be achieved according to Section 3. 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 between input sensor fault signals. The roots of the $R(s)$ polynomial matrix have been optimised for maximising the fault detection promptness, as well as to minimise the occurrence of false alarms. In order to assess the PM diagnosis technique, different fault sizes have been simulated on each sensor. Single faults in the input sensors have been generated by producing abrupt (step) and ramp (slowly developing) faults in the input signals $c(t)$. The residual signals indicate fault occurrence according to whether their values are lower or higher than the thresholds fixed in fault-free conditions. The residual processing methods can be based on simple residual geometrical analysis or comparison with fixed thresholds [3]. More complex residual evaluation can rely on statistical properties of the residuals and hypothesis testing [6], or based on adaptive thresholds, that is, the so-called threshold selector [21].

In this paper, the threshold test for FDI is performed with the logic described by (86):

$$\bar{r} - \nu\sigma_r \le r(t) \le \bar{r} + \nu\sigma_r \quad \text{for } f(t) = 0,$$
$$r(t) < \bar{r} - \nu\sigma_r \quad \text{or} \quad r(t) > \bar{r} + \nu\sigma_r \quad \text{for } f(t) \ne 0, \tag{86}$$

that is, the comparison of $r(t)$ with respect its statistical normal characteristics. $\bar{r}$ and σ_r^2 are the normal values for the mean and variance of the fault-free residual, respectively. 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 measurement errors, linearised model approximations, and disturbance signals that are not completely decoupled.

Thus, in this case, a suitable value of $\nu = 4$ for the computation of the positive and negative thresholds in (86) has been considered. To summarise the performances of the PM FDI scheme, the minimal detectable step faults on the various input sensors are collected in Table 2.

On the other hand, the minimal detectable ramp faults are reported in Table 3.

Table 2: PM FDI technique: minimal *step* faults with $\nu = 4$.

$c_i(t)$	Variable	Fault size	Delay time
Elevator	δ_e	$2°$	18 s
Aileron	δ_a	$3°$	6 s
Rudder	δ_r	$4°$	8 s
Throttle aperture %	δ_{th}	2%	15 s

Table 3: PM FDI technique: minimal *ramp* faults with $\nu = 4$.

$c_i(t)$	Variable	Fault size	Delay time
Elevator	δ_e	$0.11°/s$	26 s
Aileron	δ_a	$0.50°/s$	11 s
Rudder	δ_r	$0.49°/s$	12 s
Throttle aperture %	δ_{th}	$0.13°/s$	19 s

Table 4: NLGA FDI technique: minimal *step* faults with $\nu = 8$.

$c_i(t)$	Variable	Fault size	Delay time
Elevator	δ_e	$2°$	5 s
Aileron	δ_a	$2°$	3 s
Rudder	δ_r	$2°$	6 s
Throttle aperture %	δ_{th}	6%	3 s

Table 5: NLGA FDI technique: minimal *ramp* faults with $\nu = 8$.

$c_i(t)$	Variable	Fault size	Delay time
Elevator	δ_e	$0.21°/s$	11 s
Aileron	δ_a	$0.45°/s$	12 s
Rudder	δ_r	$0.32°/s$	10 s
Throttle aperture %	δ_{th}	$0.15°/s$	15 s

Concerning the NLGA, the synthesis of the residual generators 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 1. For example, with reference to the fourth residual generator, Procedure 1 has led to $K_{\delta_{th}} = 1$ which satisfies the norm bounds $\gamma = 1.2$ and $\beta = 400$. This guarantees a good separation of the residual signal with $\|f\|_{\mathcal{L}_2} \ge 0.05$ and $\|d\|_{\mathcal{L}_2} \le 10$, where $\mathcal{L}_2$-norm is considered.

In order to assess the NLGA diagnosis technique, single step and ramp faults have been used. Moreover, also in this case the threshold values have been chosen in simulation according to (86). A suitable value of $\nu = 8$ for the computation of the positive and negative thresholds in (86) has been considered. For what concern NLGA FDI scheme, the minimal detectable step faults on the various input sensors are summarised in Table 4.

On the other hand, the minimal detectable input sensor ramp faults are reported in Table 5.

The minimal detectable step fault values in Tables 2, 3, 4, and 5 are expressed in the unit of measure of the sensor signals. The fault step sizes and ramp slopes are relative to the
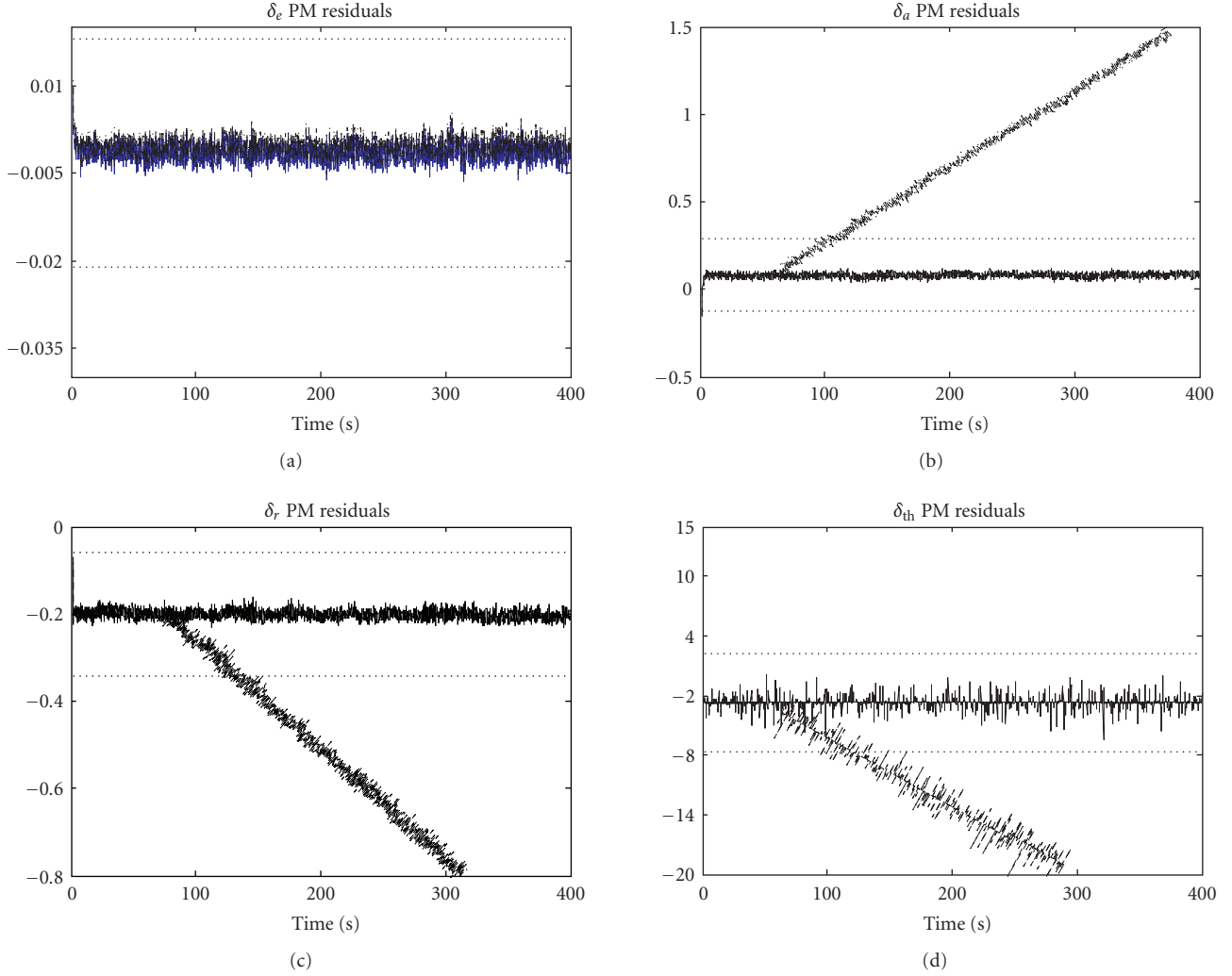
FIGURE 1: PM residuals for the 1st input sensor ramp fault $f_{c_1}(t)$ isolation with $\nu = 4$.

case in which the occurrence of a fault is detected and isolated as soon as possible. The detection delay times represent the worst case results, as they are evaluated on the basis of the time taken by the slowest residual function to cross the settled threshold. These experiments represent a further validation of the residual generator robustness with respect to the fault type, as the the residual function sensitivity was optimised only with respect to step faults.

As an example, Figure 1 shows the 4 PM residual functions generated for the complete trajectory. On the basis of the fault-free and faulty conditions, this bank provides the correct isolation of the considered input sensor *ramp* fault.

The horizontal lines represent the levels of the fault-free thresholds that are settled according to test (86) with $\nu = 4$. The first residual function depicted in Figure 1, provides also the isolation of the fault $f_c(t)$ regarding the 1st input sensor.

The second example of Figure 2 shows the 4 residual functions generated by the NLGA filter bank applied to the complete aircraft trajectory. The horizontal lines represent the thresholds with $\nu = 12$. Note that, due to the NLGA de-

sign technique, only the 1st residual related to the δ_e signal of the filter bank is sensitive to a *ramp* fault affecting the 1st input sensor.

5.1. Reliability 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 scheme [3] and the NN technique [7]. In particular, 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 an FDI system by acting as a source of false faults. The procedure recalled here requires the design of a 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
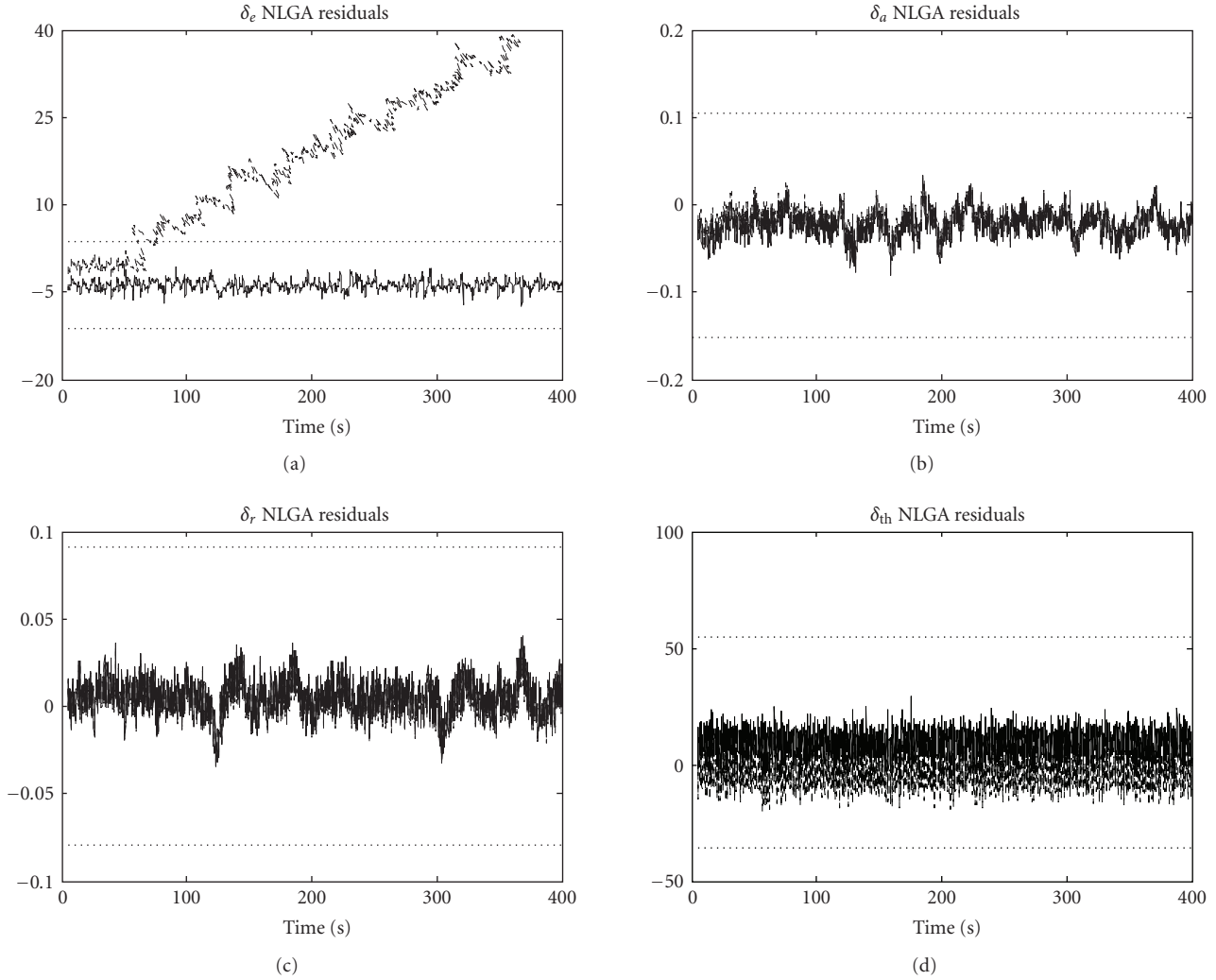
FIGURE 2: NLGA residuals for the 1st input sensor fault isolation with $\nu = 12$.

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 [3, Section 3.5, pages 99–105], whilst the noise covariance matrices were estimated as described in [22, Section 3.3, pages 70–74 and Section 4.6, pages 130-131]. Each of the 4 UIKF of the bank was decoupled from both one input sensor fault and the wind gust disturbance component, thus providing the optimal filtering of the input-output measurement noise sequences. On the other hand, 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 pro-

cess outputs. The FDI is therefore achieved by monitoring residual changes. The NN learning is typically an offline 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) [3], a bank of 4 time-delayed three-layers multilayer 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 [7].

In the following of this section, 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 3. 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
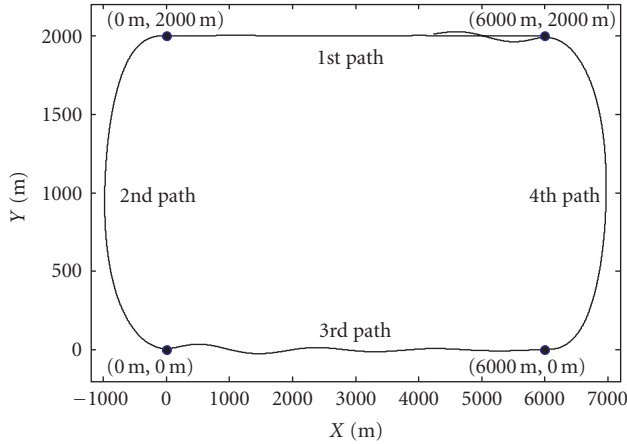
FIGURE 3: Aircraft complete trajectory example.

TABLE 6: Performances of the FDI schemes for a complete aircraft trajectory.

Variable	PM	NLGA	UIKF	NN
ν	4	12	9	5
δ_e	4°	3°	4°	3°
δ_a	5°	3°	5°	4°
δ_r	5°	3°	4°	4°
δ_{th}	7%	10%	11%	12%
Mean detection delay	26 s	25 s	31 s	27 s

alternatively the 4 branches of the reference trajectory:

(i) *straight flight condition:* true airspeed = 50 [m/s]; radius of curvature = ∞; flight-path angle = 0°; altitude = 330 [m]; flap deflection = 0°;

(ii) *turn flight condition:* true airspeed = 50 [m/s]; radius of curvature = 1000 [m]; flight-path angle = 0°; altitude = 330 [m]; flap deflection = 0°.

The reference turn flight condition is used to design the PM and the NLGA filters. The achieved results are reported in Tables 2 and 4, respectively. The performed tests represent a 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.

As an example, the fault-free and faulty residuals generated by the designed NN and UIKF banks are shown in Figures 4 and 5, respectively.

Table 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 2nd reference flight condition of the complete trajectory. Table 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 (86). The mean detection delay is also reported

in Table 6 in order to compare the effectiveness of the different FDI schemes.

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 tradeoff between two objectives, that is, 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 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 FDI scheme 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.

5.2. Monte Carlo analysis

In this section, further experiment results have been reported. They regard the performance evaluation of the developed FDI scheme 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

(a)

(b)

(c)

(d)
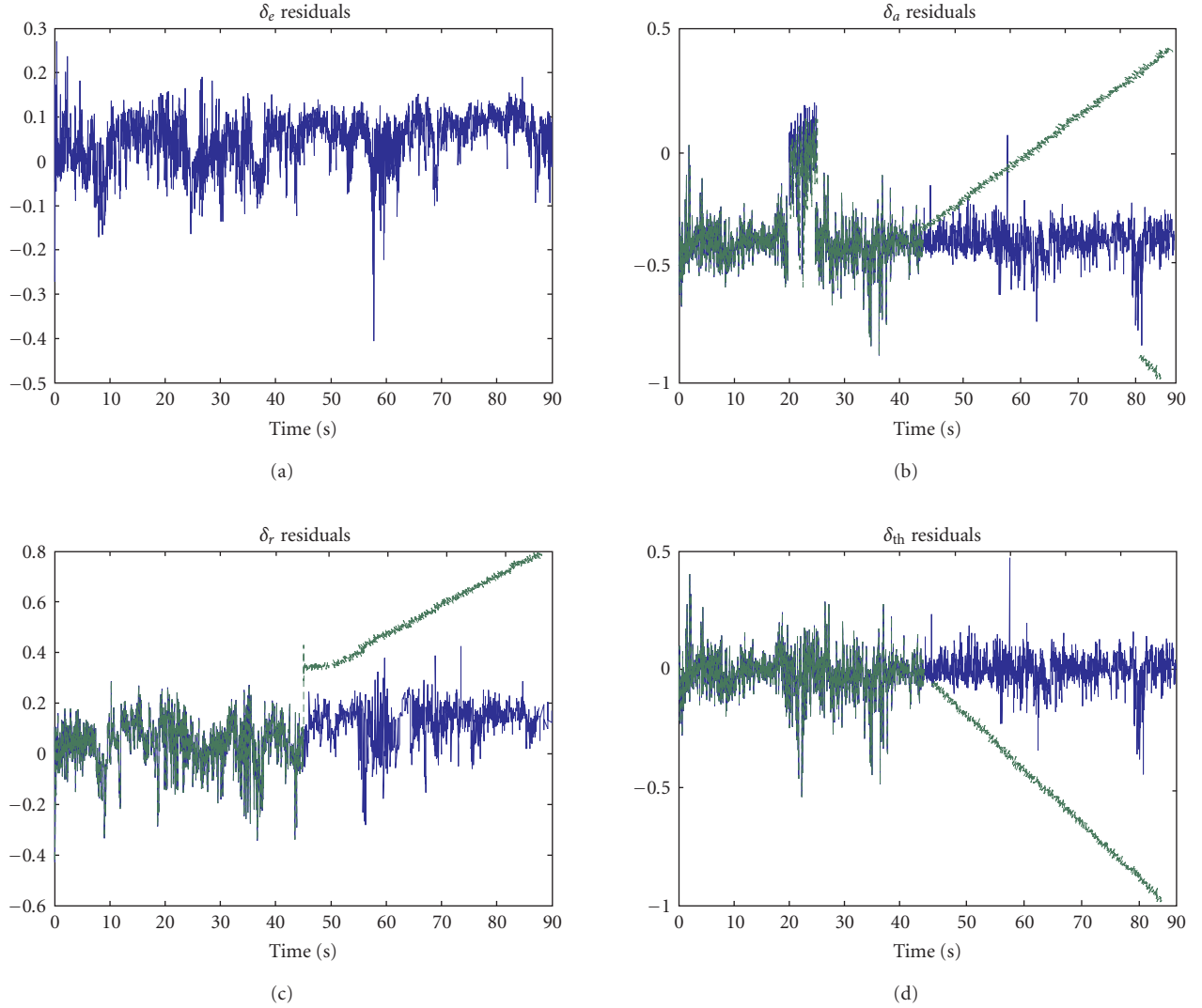
FIGURE 4: NN residuals with *ramp* fault.

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 performance 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. Table 7 summarises the results obtained by considering the PM 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 NLGA, NN, and UIKF FDI schemes. The results are summarised in Tables 8, 9, and 10.

Tables 7, 8, 9, and 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 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
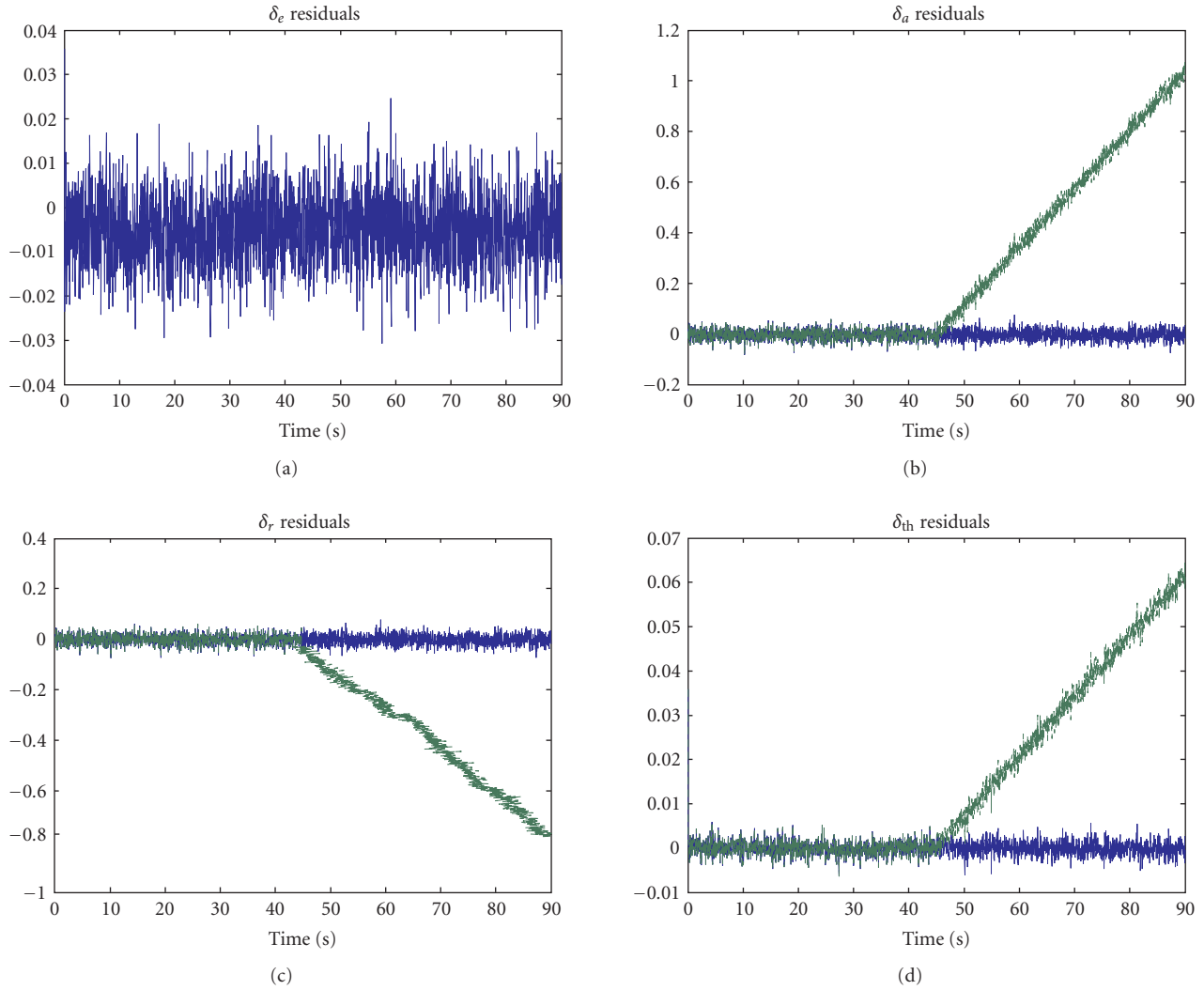
FIGURE 5: UIKF residuals with *ramp* fault.

TABLE 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 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 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

TABLE 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

uncertainty ($p = 10\%$) and fixed measurement errors. This last simulation technique example hence facilitates an assess-

ment of the reliability of the developed, analysed, and applied FDI methods.

6. CONCLUSION

The paper provided the development and application of two FDI techniques based on a PM scheme and on an NLGA method, respectively. The PM procedure led to residual generators optimising the tradeoff between disturbance-decoupling and fault sensitivity. Moreover, the application of the PM FDI scheme resulted robust with respect to model uncertainties. On the other hand, the NLGA relies on a novel design scheme based on the structural decoupling of disturbances and modelling errors, Thus, the mixed $\mathcal{H}_-/\mathcal{H}_\infty$ optimisation of the tradeoff between fault sensitivity, disturbances, and modelling errors has been proposed. The PM and NLGA residual generators were tested by considering a nonlinear aircraft simulator model that takes into account also the wind gusts, the Dryden turbulence, the input-output sensors measurement errors, as well as the engine and the servo actuators. Moreover, in order to verify the robustness characteristics and the achievable performances of the approaches, the simulation results considered a typical aircraft reference trajectory consisting of several steady-state flight conditions. The effectiveness of the developed PM and NLGA FDI schemes was shown by simulations and a comparison with widely used data-driven and model-based disturbance-decoupling FDI schemes, such as NN and UIKF diagnosis methods, was provided. The reliability and the robustness properties of the proposed residual generators to model uncertainty and disturbances and measurement noise for the aircraft nonlinear model were investigated via Monte Carlo simulations. Further works extensive comparative studies for robustness of the FDI algorithms when applied to real data.

REFERENCES

[1] A. Marcos, S. Ganguli, and G. J. Balas, "An application of H_∞ fault detection and isolation to a transport aircraft," *Control Engineering Practice*, vol. 13, no. 1, pp. 105–119, 2005.

[2] F. Amato, C. Cosentino, M. Mattei, and G. Paviglianiti, "A direct/functional redundancy scheme for fault detection and isolation on an aircraft," *Aerospace Science and Technology*, vol. 10, no. 4, pp. 338–345, 2006.

[3] J. Chen and R. J. Patton, *Robust Model-Based Fault Diagnosis for Dynamic Systems*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1999.

[4] R. Isermann, *Fault-Diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance*, Springer, Berlin, Germany, 1st edition, 2005.

[5] J. Gertler, *Fault Detection and Diagnosis in Engineering Systems*, Marcel Dekker, New York, NY, USA, 1998.

[6] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes: Theory and Application*, Prentice Hall, Upper Saddle River, NJ, USA, 1993.

[7] J. Korbicz, J. M. Koscielny, Z. Kowalczuk, and W. Cholewa, Eds., *Fault Diagnosis: Models, Artificial Intelligence, Applications*, Springer, Berlin, Germany, 1st edition, 2004.

[8] M. A. Massoumnia, *A geometric approach to failure detection and identification in linear systems*, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, Mass, USA, 1986.

[9] H. Hammouri, M. Kinnaert, and E. H. El Yaagoubi, "Observer-based approach to fault detection and isolation for nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 44, no. 10, pp. 1879–1884, 1999.

[10] C. De Persis and A. Isidori, "A geometric approach to nonlinear fault detection and isolation," *IEEE Transactions on Automatic Control*, vol. 45, no. 6, pp. 853–865, 2001.

[11] C. De Persis, R. De Santis, and A. Isidori, "Nonlinear actuator fault detection and isolation for a VTOL aircraft," in *Proceedings of the American Control Conference (ACC '01)*, vol. 6, pp. 4449–4454, Arlington, Va, USA, June 2001.

[12] M. Bonfè, P. Castaldi, W. Geri, and S. Simani, "Fault detection and isolation for on-board sensors of a general aviation aircraft," *International Journal of Adaptive Control and Signal Processing*, vol. 20, no. 8, pp. 381–408, 2006.

[13] M. Bonfè, P. Castaldi, W. Geri, and S. Simani, "Design and performance evaluation of residual genertors for the FDI of an aircraft," *International Journal of Automation and Computing*, vol. 4, no. 2, pp. 156–163, 2007.

[14] B. L. Stevens and F. L. Lewis, *Aircraft Control and Simulation*, John Wiley & Son, New York, NY, USA, 2nd edition, 2003.

[15] R. P. Guidorzi, "Invariants and canonical forms for systems stuctural and parametric identification," *Automatica*, vol. 17, no. 1, pp. 117–133, 1981.

[16] S. Van Huffel and P. Lemmerling, Eds., *Total Least Squares and Errors-in-Variables Modeling: Analysis, Algorithms and Applications*, Springer, Berlin, Germany, 1st edition, 2002.

[17] M. Bonfè, S. Simani, P. Castaldi, and W. Geri, "Residual generator computation for fault detection of a general aviation aircraft," in *Proceedings of the 16th IFAC Symposium on Automatic Control in Aerospace (ACA '04)*, vol. 2, pp. 318–323, St. Petersburg, Russia, June 2004.

[18] C. De Persis and A. Isidori, "On the observability codistributions of a nonlinear system," *Systems and Control Letters*, vol. 40, no. 5, pp. 297–304, 2000.

[19] P. Castaldi, W. Geri, M. Bonfè, and S. Simani, "Nonlinear actuator fault detection and isolation for a general aviation aircraft," in *Proceedings of the 17th IFAC Symposium on Automatic Control in Aerospace (ACA '07)*, vol. CD-Rom, pp. 1–6, Toulouse, France, June 2007.

[20] M. Hou and R. J. Patton, "LMI approach to H_-/H_∞ fault detection observers," in *Proceedings of the 1996 UKACC International Conference on Control (CONTROL '96)*, pp. 305–310, Exeter, UK, September 1996.

[21] A Emami-Naeini, M. M. Akhter, and S. M. Rock, "Effect of model uncertainty on failure detection: the threshold selector," *IEEE Transactions on Automatic Control*, vol. 33, no. 12, pp. 1106–1116, 1988.

[22] S. Simani, C. Fantuzzi, and R. J. Patton, *Model-Based Fault Diagnosis in Dynamic Systems Using Identification Techniques*, Advances in Industrial Control, Springer, London, UK, 1st edition, 2002.