ON DESIGNING LEAST ORDER RESIDUAL GENERATORS FOR FAULT DETECTION AND ISOLATION

Andras Varga

German Aerospace Center, DLR-Oberpfaffenhofen
Institute of Robotics and Mechatronics
D-82234 Wessling, Germany
Email: andras.varga@dlr.de

Abstract: This paper addresses the problem of designing residual generators with least dynamical order to solve a class of fault detection and isolation problems. The main computational ingredients are discussed for the determination of least order residual generators achieving a desired fault-to-residual influence matrix. An application of the proposed techniques to the pitch axis actuator fault monitoring for a Boeing 747 aircraft is presented.

Keywords: Fault detection, minimum degree design, dynamic covers, aircraft fault monitoring

1. INTRODUCTION

The design of fault detectors with least dynamical orders is an important aspect from practical point of view. Although there are many approaches proposed in the literature to design fault detectors, as for example (Ding and Frank, 1990; Gertler, 1998; Chen and Patton, 1999) and literature cited therein, only very few of them address explicitly the design of least order detectors (Yuan *et al.*, 1997; Frisk and Nyberg, 2001; Varga, 2003*a*). An apparently open problem is to determine a least order residual generator which achieves a desired fault-to-residual influence matrix. This problem has as main application the solution of various fault isolation problems (Patton and Hou, 1998; Gertler, 2000) and is traditionally addressed by designing a bank of single-output residual generation filters. However, the reduction of the resulting total order of the bank of filters was considered only in (Yuan *et al.*, 1997) and at present it appears that no systematic approach exists to solve this problem.

In this paper we present an enhanced approach, along the lines of dynamic covers based technique of Varga (2003*a*), to design single-output residual generation filters. A basic computational ingredient is a reliable numerical algorithm to compute least order rational nullspaces of rational matrices using state-space methods (Varga, 2003*a*). The main computation in this algorithm is the orthogonal reduction of the system pencil matrix to a Kronecker-like form, which allows to obtain, practically without any additional computation, a least order rational nullspace basis. This approach can be combined with coprime factorization techniques to determine stable rational bases which are candidate solutions to the fault detection problem. The existence conditions of the solution can be easily checked using the outcome of the nullspace computation algorithm. Least order fault detectors can be obtained by selecting an appropriate linear combination of the basis vectors by eliminating non-essential dynamics using dynamic covers based manipulations. We also discuss the least order design problem for a given fault influence specification. For this purpose, a bank of least order residual filters is determined by computing nullspace bases related to each row of the fault influence matrix. The effectiveness of this approach in conjunction with the assignment of a uniform dynamics to all detectors is illustrated via an application to the pitch axis actuator fault monitoring for a Boeing 747 100/200 aircraft.

2. DESIGN OF LEAST ORDER DETECTORS

Consider the linear time-invariant system described by the input-output relations

$$\mathbf{y}(\lambda) = G_u(\lambda)\mathbf{u}(\lambda) + G_d(\lambda)\mathbf{d}(\lambda) + G_f(\lambda)\mathbf{f}(\lambda), \quad (1)$$

where $\mathbf{y}(\lambda)$, $\mathbf{u}(\lambda)$, $\mathbf{f}(\lambda)$, and $\mathbf{d}(\lambda)$ are Laplace- or Z-transformed vectors of the p-dimensional system output vector $y(t)$, m_u-dimensional control input vector $u(t)$, m_f-dimensional fault signal vector $f(t)$, and m_d-dimensional disturbance vector $d(t)$, respectively, and where $G_u(\lambda)$, $G_f(\lambda)$ and $G_d(\lambda)$ are the *transfer-function matrices* (TFMs) from the control inputs to outputs, fault signals to outputs, and disturbances to outputs, respectively. According to the system type, $\lambda = s$ in the case of a continuous-time system or $\lambda = z$ in the case of a discrete-time system.

A residual generator for fault detection must fulfill two basic requirements: (1) to generate zero residuals in the fault-free case, for arbitrary control and disturbance inputs; (2) to generate nonzero residuals when any fault occurs in the system. These requirements can be made precise by looking for a linear residual generator (or detector) of least dynamical order having the general form

$$\mathbf{r}(\lambda) = R(\lambda)\begin{bmatrix} \mathbf{y}(\lambda) \\ \mathbf{u}(\lambda) \end{bmatrix} \quad (2)$$

such that: (i) $r(t) = 0$ when $f(t) = 0$ for all $u(t)$ and $d(t)$; and (ii) $r(t) \neq 0$ when $f_i(t) \neq 0$, for $i = 1, \ldots, m_f$. Besides the requirement that the TFM of the detector $R(\lambda)$ has least possible McMillan degree, it is also necessary, for physical realizability, that $R(\lambda)$ is a proper and stable TFM. As detector, we can always choose $R(\lambda)$ as a rational row vector.

The fulfillment of requirement (ii) ensures that faults produce non-zero residual responses. When designing fault detectors this requirement for *fault detectability* is usually replaced by the stronger request that persistent (constant) faults produce asymptotically persistent (constant) residuals. This requirement is known as *strong fault detectability* and has a special importance for practical applications.

The requirements (i) and (ii) can be easily transcribed in equivalent algebraic conditions. The condition (i) is equivalent to

$$R(\lambda)G(\lambda) = 0 \quad (3)$$

where

$$G(\lambda) = \begin{bmatrix} G_u(\lambda) & G_d(\lambda) \\ I_{m_u} & 0 \end{bmatrix}, \quad (4)$$

while the detectability condition (ii) is equivalent to

$$R_f^{(i)}(\lambda) \neq 0, \quad i = 1, \ldots, m_f \quad (5)$$

where $R_f^{(i)}(\lambda)$ is the i-th column of

$$R_f(\lambda) := R(\lambda)\begin{bmatrix} G_f(\lambda) \\ 0 \end{bmatrix} \quad (6)$$

Enforcing the strong detectability of constant faults is equivalent to ensure finite non-zero DC-gains for each column of $R_f(\lambda)$, thus to ask

$$0 < \left\| R_f^{(i)}(\lambda_s) \right\| < \infty, \quad i = 1, \ldots, m_f \quad (7)$$

where $\lambda_s = 0$ or $\lambda_s = 1$ depending on the system type (continuous or discrete).

From (3) it appears that $R(\lambda)$ is a left annihilator of $G(\lambda)$, thus one possibility to determine $R(\lambda)$ is to compute first a minimal basis $N_l(\lambda)$ for the *left nullspace* of $G(\lambda)$, and then to build a stable scalar output detector as

$$R(\lambda) = h(\lambda)N_l(\lambda), \quad (8)$$

representing a linear combination of the rows of $N_l(\lambda)$, such that conditions (5) or (7) are fulfilled. The above expression of $R(\lambda)$ represents a parametrization of *all* possible detectors and is the basis for the class of *nullspace methods* introduced in (Frisk and Nyberg, 2001). While this work relies on using polynomial nullspace bases for $N_l(\lambda)$, an alternative approach relying on proper rational bases has been proposed by the author (Varga, 2003*a*). The main advantage of this latter method is to rely exclusively on reliable numerical techniques based on state-space computations.

In what follows, we discuss more in details the two main design steps, namely, the determination of $N_l(\lambda)$ and the choice of appropriate $h(\lambda)$ and we discuss reliable numerical methods to perform these steps.

3. COMPUTATION OF LEAST ORDER RATIONAL NULLSPACE BASES

The approach to compute least order detectors proposed in this paper is strongly connected to the details of the computations of the left nullspace basis $N_l(\lambda)$. Therefore, in this section we summarize the computational approach of (Varga, 2003*a*) to determine least order rational nullspace bases of a rational matrix which underlies the proposed procedure. This method is also equivalent to the pencil approach to design fault detectors proposed in (Patton and Hou, 1998).

We assume the system (1) has the following *descriptor* state space realization

$$\begin{aligned} E\lambda x(t) &= Ax(t) + B_u u(t) + B_d d(t) + B_f f(t) \\ y(t) &= Cx(t) + D_u u(t) + D_d d(t) + D_f f(t) \end{aligned} \quad (9)$$

with the n-dimensional state vector $x(t)$ and where $\lambda x(t) = \dot{x}(t)$ or $\lambda x(t) = x(t+1)$ depending on the type of the system, continuous or discrete, respectively. In general, the square matrix E can be singular, but we will assume that the linear pencil $A - \lambda E$ is regular. For proper TFMs in (1), we can choose a *standard* state space realization where $E = I$. In general we can assume that the representation (9) is minimal, that is, the pair $(A - \lambda E, C)$ is *observable* and the pair $(A - \lambda E, [\, B_u \; B_d \; B_f \,])$ is *controllable*.

The corresponding TFMs of the model in (1) are

$$\begin{aligned} G_u(\lambda) &= C(\lambda E - A)^{-1} B_u + D_u \\ G_d(\lambda) &= C(\lambda E - A)^{-1} B_d + D_d \\ G_f(\lambda) &= C(\lambda E - A)^{-1} B_f + D_f \end{aligned}$$

The $(p + m_u) \times (m_u + m_d)$ TFM $G(\lambda)$ in (4) can be realized in state space form as

$$G(\lambda) = \left[\begin{array}{c|cc} A - \lambda E & B_u & B_d \\ \hline C & D_u & D_d \\ 0 & I_{m_u} & 0 \end{array} \right]$$

Assume that the rational matrix $G_d(\lambda)$ has rank $r_d \le \min(p, m_d)$. It follows that a left nullspace basis $N_l(\lambda)$ satisfying

$$N_l(\lambda) G(\lambda) = 0 \quad (10)$$

is a $(p - r_d) \times (p + m_u)$ rational matrix whose rows are rational vectors which form a basis for the left nullspace of $G(\lambda)$.

The method described in (Varga, 2003*a*) exploits the simple fact that $N_l(\lambda)$ is a left nullspace basis of $G(\lambda)$ iff $[\, M_l(\lambda) \; N_l(\lambda) \,]$ is a left nullspace basis of the system matrix

$$S(\lambda) = \begin{bmatrix} A - \lambda E & B_u & B_d \\ C & D_u & D_d \\ 0 & I_{m_u} & 0 \end{bmatrix}.$$

Thus to compute $N_l(\lambda)$ we can determine equivalently a left nullspace basis $Y_l(\lambda)$ for $S(\lambda)$ and then $N_l(\lambda)$ simply results as

$$N_l(\lambda) = Y_l(\lambda) \begin{bmatrix} 0 \\ I_{p+m_u} \end{bmatrix}.$$

$N_l(\lambda)$ can be computed (see bellow) by employing linear pencil reduction algorithms based on orthogonal transformations. Bases with special properties (e.g., stable) can be obtained by post-processing $N_l(\lambda)$, using the fact that if $N_l(\lambda) = \widetilde{M}_l^{-1}(\lambda)\widetilde{N}_l(\lambda)$ is a stable *left coprime factorization* (LCF), then the numerator matrix $\widetilde{N}_l(\lambda)$ can be equally employed as a left nullspace basis of $G(\lambda)$.

Let Q and Z be orthogonal matrices (for instance, determined by using the algorithms of (Beelen, 1987; Varga, 1996) such that the transformed pencil $\widetilde{S}(\lambda) := QS(\lambda)Z$ is in the Kronecker-like staircase form

$$\widetilde{S}(\lambda) = \left[\begin{array}{cc} A_{r,reg} - \lambda E_{r,reg} & A_{r,l} - \lambda E_{r,l} \\ \hline 0 & A_l - \lambda E_l \\ \hline 0 & C_l \end{array} \right] \quad (11)$$

where the descriptor pair $(A_l - \lambda E_l, C_l)$ is observable, E_l is non-singular, and $A_{r,reg} - \lambda E_{r,reg}$ has full row rank excepting possibly a finite set of values of λ (i.e, the invariant zeros of $S(\lambda)$). It follows that we can choose the nullspace $\widetilde{Y}_l(\lambda)$ of $\widetilde{S}(\lambda)$ in the form

$$\widetilde{Y}_l(\lambda) = [\; 0 \mid C_l(\lambda E_l - A_l)^{-1} \mid I \;].$$

Then the nullspace of $G(\lambda)$ is

$$N_l(\lambda) = \widetilde{Y}_l(\lambda) Q \begin{bmatrix} 0 \\ I_{p+m_u} \end{bmatrix}$$

and if we partition

$$Q \begin{bmatrix} 0 \\ I_{p+m_u} \end{bmatrix} = \begin{bmatrix} B_{r,l} \\ B_l \\ D_l \end{bmatrix}$$

in accordance with the column partitioning of $\widetilde{Y}_l(\lambda)$, we obtain

$$N_l(\lambda) = C_l(\lambda E_l - A_l)^{-1} B_l + D_l \quad (12)$$

Thus, $(A_l - \lambda E_l, B_l, C_l, D_l)$, with E_l nonsingular, is a descriptor system representation for $N_l(\lambda)$. Note that, to obtain this nullspace basis, we performed exclusively orthogonal transformations on the system matrices. We can prove that all computed matrices are exact for a slightly perturbed original system. It follows that the algorithm to compute the nullspace basis is *numerically backward stable*.

Consider now the detailed structure of the full column rank subpencil $\begin{bmatrix} A_l - \lambda E_l \\ C_l \end{bmatrix}$. We can assume that this subpencil is in the following observability staircase form

$$\left[\begin{array}{c|c|c|c} A_{\ell,\ell+1} & A_{\ell,\ell} - \lambda E_{\ell,\ell} & \cdots & A_{\ell,1} - \lambda E_{\ell,1} \\ \hline & A_{\ell-1,\ell} & \ddots & \vdots \\ \hline & & \ddots & A_{1,1} - \lambda E_{1,1} \\ \hline & & & A_{0,1} \end{array} \right] \quad (13)$$

where $A_{i,i+1} \in \mathbb{R}^{\mu_i \times \mu_{i+1}}$, with $\mu_{\ell+1} = 0$, are full column rank upper triangular matrices, for $i = 0, \ldots, \ell$. Note that this form is automatically obtained by using the pencil reduction algorithms described in (Beelen, 1987; Varga, 1996). The left (or row) Kronecker indices of $G(\lambda)$ result as follows: there are $\mu_{i-1}-\mu_i$ Kronecker blocks of size $i\times(i-1)$, for $i=1,\ldots,\ell+1$. The row dimension of $N_l(\lambda)$ is given by the total number of Kronecker indices, thus

$$\sum_{i=1}^{\ell+1}(\mu_{i-1}-\mu_i)=\mu_0$$

It follows that $\mu_0 := p-r_d$. The order of the realization (12) is $n_l := \sum_{i=1}^{\ell}\mu_i$ and it was shown in (Varga, 2003*a*) that this order is equal to the least possible one.

4. CHECKING FAULT DETECTABILITY

We discuss shortly how conditions (5) or (7) can be efficiently checked for a candidate detector corresponding to the computed left nullspace basis $N_l(\lambda)$. For this, we use the outcome of the left nullspace algorithm to compute

$$Q\begin{bmatrix}B_f\\ D_f\\ 0\end{bmatrix}=\begin{bmatrix}*\\ \hline \widetilde{B}_f\\ \hline \widetilde{D}_f\end{bmatrix},$$

where the row partitioning of the right hand side corresponds to the column partitioning of $\widetilde{Y}_l(\lambda)$. It easy to show that

$$R_f(\lambda)=\widetilde{Y}_l(\lambda)Q\begin{bmatrix}B_f\\ D_f\\ 0\end{bmatrix}=C_l(\lambda E_l-A_l)^{-1}\widetilde{B}_f+\widetilde{D}_f$$

Since the pair $(A_l-\lambda E_l, C_l)$ is observable, checking the condition (5) is equivalent to verify that

$$\begin{bmatrix}\widetilde{B}_f^{(i)}\\ \widetilde{D}_f^{(i)}\end{bmatrix}\neq 0,\quad i=1,\ldots,m_f$$

where $\widetilde{B}_f^{(i)}$ and $\widetilde{D}_f^{(i)}$ denote the i-th columns of $\widetilde{B}_f$ and $\widetilde{D}_f$, respectively.

To check the strong detectability condition (7), λ_s must not be a zero of $R_f^{(i)}(\lambda)$. Direct evaluation of $R_f^{(i)}(\lambda_s)$ may be not possible if $N_l(\lambda)$, and thus presumably also $R_f^{(i)}(\lambda)$, has poles in λ_s. However, such poles can be easily perturbed by using a randomly generated output injection K, to work, instead of $R_f(\lambda)$, with the perturbed

$$\widetilde{R}_f(\lambda)=C_l(\lambda E_l-A_l-KC_l)^{-1}(\widetilde{B}_f+K\widetilde{D}_f)+\widetilde{D}_f$$

5. DETECTOR ORDER REDUCTION

Let $N_l(\lambda)$ be a left proper nullspace basis of $G(\lambda)$ (i.e., satisfying $N_l(\lambda)G(\lambda)=0$). In this section we will address the choice of a vector $h(\lambda)$ such that the scalar output detector $R(\lambda) := h(\lambda)N_l(\lambda)$ has least McMillan degree.

Assume $N_l(\lambda)$ has the standard state space realization

$$N_l(\lambda)=\left[\begin{array}{c|c}A_l-\lambda I_{n_l} & B_l\\ \hline C_l & D_l\end{array}\right] \qquad (14)$$

where we replaced A_l and B_l by $E_l^{-1}A_l$ and $E_l^{-1}B_l$, respectively. It is easy to show that instead $N_l(\lambda)$ we can equally use

$$\widetilde{N}_l(\lambda)=\left[\begin{array}{c|c}A_l+KC_l-\lambda I & B_l+KD_l\\ \hline C_l & D_l\end{array}\right]$$

where K is an arbitrary output injection matrix. Moreover, all solutions of order at most n_l can be generated in the form

$$R(\lambda)=\left[\begin{array}{c|c}A_l+KC_l-\lambda I_{n_l} & B_l+KD_l\\ \hline hC_l & hD_l\end{array}\right] \qquad (15)$$

where h is chosen such that (5) (or (7)) is fulfilled.

Thus the least order design problem can be reformulated as follows: for a given h determine the injection matrix K such that a maximum number of eigenvalues of A_l+KC_l becomes unobservable for the pair (A_l+KC_l, hC_l). In a dual setting in terms of transposed matrices, the least order solution of the above problem can be recast as a minimal order dynamic cover design problem.

Consider the set

$$\mathcal{J}=\{\mathcal{V} : A_l^T\mathcal{V}\subset \operatorname{Im} C_l^T+\mathcal{V},\quad \operatorname{Im}\left(C_l^T h^T\right)\subset\mathcal{V}\}$$

and let $\mathcal{J}^*$ denote the set of subspaces $\mathcal{J}$ of least dimension. If $\mathcal{V}\in\mathcal{J}^*$ then a matrix K^T can be determined such that

$$\left(A_l^T+C_l^TK^T\right)\mathcal{V}+\operatorname{Im}\left(C_l^Th^T\right)\subset\mathcal{V}$$

Thus, determining a minimal dimension subspace $\mathcal{V}$ is equivalent to a Type I minimal order cover design problem, and a computational approach to solve it has been proposed in (Varga, 2003*b*). The outcome of this method is, besides $\mathcal{V}$, the matrix K which achieves a maximal order reduction by pole-zero cancellations.

The details of this procedure are presented in (Varga, 2003*b*). The general idea is to perform a preliminary orthogonal similarity transformation on the system matrices in (14) by applying a special version of the controllability staircase form algorithm (see for example (Varga, 1981)) to the pair $\left([\,C_l^Th^T\ C_l^T\,], A_l^T\right)$. With additional block permutations and block row/column transformations, we can bring the transformed system matrices in the following *special form*

$$V^{-1}A_lV=\left[\begin{array}{c|c}\widehat{A}_{11} & \widehat{A}_{12}\\ \hline \widehat{A}_{21} & \widehat{A}_{22}\end{array}\right],\quad V^{-1}B_l=\left[\begin{array}{c}\widehat{B}_1\\ \hline \widehat{B}_2\end{array}\right],$$

$$\left[\begin{array}{c}C_l\\ \hline hC_l\end{array}\right]V=\left[\begin{array}{c|c}\widehat{C}_{11} & \widehat{C}_{12}\\ \hline 0 & \widehat{C}_{22}\end{array}\right],$$

where the pairs $(\widehat{C}_{11},\widehat{A}_{11})$ and $(\widehat{C}_{22},\widehat{A}_{22})$ are observable, and the submatrices $\widehat{C}_{11}$ and $\widehat{A}_{21}$ have the particular structure

$$\begin{bmatrix} \widehat{A}_{21} \\ \widehat{C}_{11} \end{bmatrix} = \begin{bmatrix} 0 & A_{21} \\ 0 & C_{11} \end{bmatrix}$$

with C_{11} having full column rank. Thus, by taking

$$K = V \begin{bmatrix} 0 \\ K_2 \end{bmatrix}$$

with K_2 satisfying $K_2 C_{11} + A_{21} = 0$, we achieve the cancellation of the maximum number of unobservable eigenvalues. The resulting detector of least McMillan degree, obtained by deleting the unobservable part, is

$$\widetilde{R}(\lambda) = \left[\begin{array}{c|c} \widehat{A}_{22} + K_2 \widehat{C}_{12} - \lambda I & \widehat{B}_2 + K_2 D_l \\ \hline \widehat{C}_{22} & h D_l \end{array} \right]$$

The main aspect which remains to be discussed is the choice of matrix h. Generally with a randomly generated h one achieves a detector whose order corresponds to the maximum degree of a minimal polynomial basis. This result can be justified by looking at the details of determining polynomial basis vectors. Using the Kronecker-like form based method to determine a minimal polynomial basis $N_l(\lambda)$ proposed by Beelen (1987) (see also (Varga, 2003*a*)), the left polynomial basis is determined using the resulting staircase form of the observable pair $(A_l - \lambda E_l, C_l)$ in (13). The left Kronecker indices μ_i, $i = 0, 1, \ldots, \ell+1$, (recall that $\mu_{\ell+1} = 0$) provide the complete information on the number of polynomial basis vectors of a given degree. The main result of (Beelen, 1987, Section 4.6) in this respect states that there are $\mu_{i-1} - \mu_i$ polynomial vectors of degree $i - 1$. The maximum degree is thus ℓ and this is precisely the order which can be achieved applying the cover algorithm. In the case when no disturbance inputs are present, this is a well know result in designing functional observers (Luenberger, 1966).

Lower orders detectors can be obtained using particular choices of the row vector h. By exploiting the details of the staircase form (13) of the pair $(A_l - \lambda E_l, C_l)$ (see (Beelen, 1987)) we can show that if only the last j elements of h are nonzero (e.g., chosen randomly), then the corresponding least order solution corresponds to a combination of the first j vectors of the polynomial basis. In this way, a systematic search can be performed by generating successive choices of h with increasing number of nonzero elements and checking for the resulting residual generator the conditions (5) (or (7)). The resulting detectors have non-decreasing orders and thus the first detector satisfying these conditions represents a satisfactory least order design. For more details see (Varga, 2007).

6. RELATION TO OTHER APPROACHES

We discuss shortly the relationships of the proposed method with other methods proposed in the literature. The method of (Yuan *et al.*, 1997) achieves the order reduction of the detector by assigning the eigenstructure of the matrix $A_l + KC_l$ with multiple eigenvalues in a maximum number of Jordan blocks. The dimensions of these blocks are constrained by the observability indices of the pair (A_l, C_l) by the well-known fundamental theorem of linear state variable feedback due to Rosenbrock (1970, p.190). In single-output only one Jordan block can exist for each distinct eigenvalue, and therefore this technique can be used to cancel a maximum number of unobservable eigenvalues. However, the computational procedure sketched in (Yuan *et al.*, 1997) relies on a delicate eigenstructure assignment and is not numerically reliable. Note that assigning multiple eigenvalues leads in general to an increased sensitivity of the eigenvalue problem. In a second computational step, the non-observable part of the resulting system must be eliminated. Performing first an eigenstructure assignment to produce unobservable multiple eigenvalues appears to be just a computational trick and thus an unnecessary computational detour. In contrast, the method based on dynamic covers is a direct technique and involves only reliable numerical computations.

A second approach addressing the least order design has been proposed by Frisk and Nyberg (2001) and relies on a polynomial approach. The nullspace $N_l(\lambda)$ is determined as a polynomial matrix and a least order detector is determined by building linear combinations of several basis vectors. Thus, the least order results as the largest degree of the involved polynomials. Although this method is not generally recommended as a computational method due to the intrinsic ill-conditioning of polynomial representations (see (Varga, 2003*a*)) especially for higher order systems, still the approach provides insight into the structure of the problem and allows to guide the choice of least order candidates. Additional difficulties are expected when converting the polynomial basis into a rational basis and checking conditions (5) (or (7)).

The last method we mention was proposed by the author in (Varga, 2003*a*). The order reduction can be achieved by selecting two disjoint components $N_{l,1}(\lambda)$ and $N_{l,2}(\lambda)$ of the nullspace $N_l(\lambda)$ and computing appropriate $Y(\lambda)$ such that $N_{l,1}(\lambda) + Y(\lambda)N_{l,2}(\lambda)$ has least order. The underlying computational approach relies on employing Type II dynamic covers. For each design, the conditions (5) must be checked. The main problem with this approach is that there are cases when the least order can not be directly achieved. A typical case is when a linear combination of the full nullspace basis $N_l(\lambda)$ is necessary.

7. SOLVING FAULT ISOLATION PROBLEMS

The more advanced functionality of fault isolation (i.e., exact location of faults) can be often achieved by designing a bank of fault detectors (Gertler, 1998) or by direct design of fault isolation filters (Varga, 2004*b*). Designing detectors which are sensitive to some fault and insensitive to others can be reformulated as a standard fault detection problem, by formally redefining the faults to be rejected in the residual as fictive disturbances.

For a given detector $R(\lambda)$ the corresponding fault influence matrix S can be coded as

$$\begin{aligned} S_{ij} &= 1 \quad \text{if } R_{f,ij}(\lambda_s) \neq 0 \\ S_{ij} &= -1 \text{ if } R_{f,ij}(\lambda_s) = 0 \text{ but } R_{f,ij}(\lambda) \neq 0 \\ S_{ij} &= 0 \quad \text{if } R_{f,ij}(\lambda) = 0 \end{aligned}$$

If $S_{ij} = 1$ then we say that fault j is *strongly* detected in residual i. If $S_{ij} = -1$ then the fault j is only *weekly* detected in residual i. The fault j is not detected in residual i if $S_{ij} = 0$.

To solve a fault isolation problem, a given fault influence matrix must be achieved by using a bank of single output detectors. Each single output detector achieves an influence matrix representing a single row of the desired fault influence matrix. For example, to achieve the complete isolation of maximum k simultaneous faults the choice $S = I_k$ is necessary. In many practical applications such a specification can not be achieved due to the lack of sufficient number of measurements. Provided we can assume that the faults occur one at time, a so-called week isolation of k faults is possible using a specification matrix whose i-th row contains all ones excepting the element in column i which is zero. For example, for 3 faults S is chosen as

$$S = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

If this fault influence specification can be achieved, then the occurrence of fault i can be detected if all residuals (excepting the i-th residual) are non-zero. More insight on how to specify fault influence matrices can be found in (Gertler, 2000).

For an efficient implementation and operation of a residual generator (or a bank of such devices), it is generally desirable to keep the order of the fault detector(s) as low as possible. Thus, minimizing the total order of the detector achieving a given specification is important for practical applications. Determining a detector of least order for a given specification S is to the best of our knowledge still an open problem. The only paper addressing this problem is apparently (Yuan *et al.*, 1997). Unfortunately, no systematic approach can be derived on basis of the results of (Yuan *et al.*, 1997) and the possible approaches are rather ad-hoc and problem specific. In what follows we suggest an alternative approach which usually leads to detectors of quite low orders. Although this approach does not provide a solution to the least order design problem, its main advantage lies in that all computational aspects can be addressed using numerically reliable techniques.

The proposed approach is very simple and relies on the least order design techniques developed in this paper. For each row of S a single output detector of least order can be determined, leading to a global detector whose order is the sum of the orders of the component detectors. For practical reasons it is reasonable to assign approximately the same dynamics to all detectors, to ensure that the minimum detection times of all faults are almost the same. Assigning identical dynamics to all detectors has the following interesting consequence: the resulting global dynamics is usually non-minimal and thus an order reduction can be achieved by eliminating the uncontrollable/unobservable eigenvalues.

This simple trick works well in many cases leading to detectors of surprisingly low orders. We used this technique for the example of Yuan *et al.* (1997, Table 2), where a 18×9 fault influence matrix S served as specification. Each line of S can be realized by a detector of order 1 or 2 with eigenvalues $\{-1\}$ or $\{-1, -2\}$. The total order of the resulting bank of detectors is 32, but after performing the minimal realization procedure a detector of total order 6 has been obtained which achieves the same specification matrix. Recall that the detector computed in (Yuan *et al.*, 1997) has order 14.

8. MONITORING PITCH AXIS ACTUATOR FAILURES FOR A BOEING 747

A linearized nominal longitudinal model of a Boeing 747 aircraft has the form (9), where the state, input and output variables are defined as follows:

$$x = \begin{bmatrix} \delta q \\ \delta V_{TAS} \\ \delta \alpha \\ \delta \theta \\ \delta h_e \end{bmatrix} \left(\begin{array}{r} \textit{pitch rate } [rad/s] \\ \textit{true airspeed } [m/s] \\ \textit{angle of attack } [rad] \\ \textit{pitch angle } [rad] \\ \textit{altitude } [m] \end{array} \right),$$

$$u = \begin{bmatrix} \delta_{eir} \\ \delta_{eil} \\ \delta_{eor} \\ \delta_{eol} \\ \delta_{ih} \\ \delta EPR_1 \\ \delta EPR_2 \\ \delta EPR_3 \\ \delta EPR_4 \end{bmatrix} \left(\begin{array}{r} \textit{right inner elevator} [rad] \\ \textit{left inner elevator} [rad] \\ \textit{right outer elevator} [rad] \\ \textit{left outer elevator} [rad] \\ \textit{stabilizer trim angle} [rad] \\ \textit{thrust engine \#1} [rad] \\ \textit{thrust engine \#2} [rad] \\ \textit{thrust engine \#3} [rad] \\ \textit{thrust engine \#4} [rad] \end{array} \right),$$

$$y = \begin{bmatrix} \delta\alpha \\ \delta\dot{V}_{TAS} \\ \delta\theta \\ \delta q \\ \delta V_z \\ \delta h_e \end{bmatrix} \quad \begin{pmatrix} angle\ of\ attack\ [rad] \\ acceleration\ [m/s^2] \\ pitch\ angle\ [rad] \\ pitch\ rate\ [rad/s] \\ vertical\ velocity\ [m/s] \\ altitude\ [m] \end{pmatrix}$$

There is no the disturbance input $d(t)$ in this model. The state space model matrices are given in Appendix A and correspond to a given set of parameter values (mass = 317,000 kg, center of gravity coordinates: $X_{cg} = 25\%$, $Y_{cg} = 0$, $Z_{cg} = 0$) and a specific flight condition – straight-and-level flight at altitude 600 m and speed of 92.6 m/s with a flap setting at 20^o and landing gear up. For more details on the employed model and for additional references see (Varga, 2007).

The fault inputs are defined as

$$f = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \end{bmatrix} \quad \begin{pmatrix} right\ inner\ elevator\ fault[rad] \\ left\ inner\ elevator\ fault[rad] \\ right\ outer\ elevator\ fault[rad] \\ left\ outer\ elevator\ fault[rad] \\ stabilizer\ fault[rad] \end{pmatrix}$$

and thus $B_f = B_u(:, 1:5)$ and $D_f = D_u(:, 1:5)$.

The maximally achievable fault influence structure is

$$S = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

From the last three lines of S it can be observed that the isolation of faults grouped in three groups (f_1, f_2), (f_3, f_4) and f_5 is achievable, although all faults are only weekly detectable.

The simplest fault detection task is to determine if any actuator fault in the pitch axis occurred. This comes down to design a fault detector achieving the trivial influence structure

$$S_1 = \begin{bmatrix} 1\ 1\ 1\ 1\ 1 \end{bmatrix}$$

by using the lowest order dynamics. To design such a detector, the function `fd` from the Fault Detection Toolbox described in (Varga, 2006) has been used. Using the least order design option implemented using the proposed approach, a first order residual generator can be used for this task.

Provided we assume that the groups of faults (f_1, f_2), (f_3, f_4) and f_5 do not simultaneously occur, the achievable specification

$$S_2 = \begin{bmatrix} 0\ 0\ 1\ 1\ 1 \\ 1\ 1\ 0\ 0\ 1 \\ 1\ 1\ 1\ 1\ 0 \end{bmatrix}$$

can be used for week isolation using the following decision logic:

- inner elevator fault occurred if $r_1 = 0$, $r_2 \neq 0$, and $r_3 \neq 0$;
- outer elevator fault occurred if $r_1 \neq 0$, $r_2 = 0$, and $r_3 \neq 0$;
- stabilizer fault occurred if $r_1 \neq 0$, $r_2 \neq 0$, and $r_3 = 0$.

Using the least order design option, three first order detectors can be obtained using the function `fdbank` leading to a detector of total order 3. Note that without the least order design option, a detector of total order 10 results, while using the standard observer based approach (Patton and Hou, 1998), a detector of total order 15 is to be expected.

A more realistic setting is to add actuator dynamics to each input actuator-surface channel. The elevator dynamics can be approximated by transfer functions of the form $37/(s+37)$, while for the stabilizer dynamics we can take $0.5/(s+0.5)$. The resulting model has now order 10 and we can achieve the same influence structure with a bank of three detectors of total order 6.

Further enhancement of fault isolation is possible by employing direct measurements of surface positions. For example, with a single additional measurement of the stabilizer surface angle it is possible to achieve the specification

$$S_3 = \begin{bmatrix} 1\ 1\ 0\ 0\ 0 \\ 0\ 0\ 1\ 1\ 0 \\ 0\ 0\ 0\ 0\ 1 \end{bmatrix}$$

and thus to isolate inner elevator, outer elevator and stabilizer faults. The above specification can be achieved using a bank of three detectors of total order 5.

Finally, a complete fault isolation (i.e., $S_4 = I_5$) can be achieved by adding a minimal number of three surface angle measurements from the two left elevators and the stabilizer. The resulting bank of five detectors has a total order of 6.

9. CONCLUSION

We proposed a numerically sound computational approach to design least order fault detectors. The least order detector is obtained using a systematic search over single output detectors of increasing orders built as linear combinations of rational nullspace basis vectors. The order reduction is achieved using Type I minimal dynamic covers for which reliable computational techniques are available (Varga, 2003*b*). The proposed method can be effective also in obtaining low order detectors which achieve given fault influence matrices. A practical actuator fault monitoring example illustrates the potential of this approach to solve fault detection and isolation problems for a Boeing

747 aircraft. For the proposed new method software tools have been implemented in the recently developed Fault Detection Toolbox for MATLAB (Varga, 2006).

The proposed least order design method is particularly effective in determining least order residual generation filter specifications achieving given fault-to-residual influence matrices. These specifications can be then used for more realistic designs, where robustness against noisy inputs (e.g., wind gusts) and noisy measurements, as well as robustness against parametric uncertainties must achieved. The robustness aspects in conjunction with least order design will be addressed in a separate work.

REFERENCES

Beelen, Th.G.J. (1987). New Algorithms for Computing the Kronecker structure of a Pencil with Applications to Systems and Control Theory. Ph. D. Thesis. Eindhoven University of Technology.

Chen, J. and R. J. Patton (1999). *Robust Model-Based Fault Diagnosis for Dynamic Systems.* Kluwer Academic Publishers, London.

Ding, X. and P. M. Frank (1990). Fault detection via factorization. *Systems & Control Lett.* **14**, 431–436.

Frisk, E. and M. Nyberg (2001). A minimal polynomial basis solution to residual generation for fault diagnosis in linear systems. *Automatica* **37**, 1417–1424.

Gertler, J. (1998). *Fault Detection and Diagnosis in Engineering Systems.* Marcel Dekker. New York.

Gertler, J. (2000). Designing dynamic consistency relation for fault detection and isolation. *Int. J. Control* **73**, 720–732.

Luenberger, D. G. (1966). Observers for multivariable systems. *IEEE Trans. Automat. Control* **11**, 190 – 197.

Patton, R. J. and M. Hou (1998). Design of fault detection and isolation observers: a matrix pencil approach. *Automatica* **34**(9), 1135–1140.

Rosenbrock, H. H. (1970). *State-Space and Multivariable Theory.* Wiley, New York.

Varga, A. (1981). Numerically stable algorithm for standard controllability form determination. *Electron. Lett.* **17**, 74–75.

Varga, A. (1996). Computation of Kronecker-like forms of a system pencil: Applications, algorithms and software. *Proc. CACSD'96 Symposium, Dearborn, MI.* pp. 77–82.

Varga, A. (1998). Computation of coprime factorizations of rational matrices. *Lin. Alg. & Appl.* **271**, 83–115.

Varga, A. (2003*a*). On computing least order fault detectors using rational nullspace bases. *Proc. of IFAC Symp. SAFEPROCESS'2003, Washington D.C.*

Varga, A. (2003*b*). Reliable algorithms for computing minimal dynamic covers. *Proc. of CDC'2003, Maui, Hawaii.*

Varga, A. (2004*a*). Computation of least order solutions of linear rational equations. *Proc. of MTNS'04, Leuven, Belgium.*

Varga, A. (2004*b*). New computational approach for the design of fault detection and isolation filters. In: *Advances in Automatic Control* (M. Voicu, Ed.). Vol. 754, *The Kluwer International Series in Engineering and Computer Science.* Kluwer Academic Publishers. pp. 367–381.

Varga, A. (2006). A fault detection toolbox for MATLAB. *Proc. of CACSD'06, Munich, Germany.*

Varga, A. (2007). Design of least order residual generators for fault detection and isolation with application to monitoring actuator/surface faults for a Boeing 747 100/200 aircraft, Technical report, GARTEUR AG-16 TP-2007-1, DLR Oberpfaffenhofen, Inst. of Robotics and Mechatronics.

Yuan, Z., G. C. Vansteenkiste and C. Y. Wen (1997). Improving the observer-based FDI design for efficient fault isolation. *Int. J. Control* **68**(1), 197–218.

Appendix A. LINEARIZED BOEING 747 MODEL

$$A = \begin{bmatrix} -0.4861 & 0.000317 & -0.5588 & 0 & -2.04\cdot 10^{-6} \\ 0 & -0.0199 & 3.0796 & -9.8048 & 8.98\cdot 10^{-5} \\ 1.0053 & -0.0021 & -0.5211 & 0 & 9.30\cdot 10^{-6} \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -92.6 & 92.6 & 0 \end{bmatrix}$$

$$B_u = \begin{bmatrix} -0.1455 & -0.1455 & -0.1494 & -0.1494 & -1.2860 & 0.0013 & 0.0035 & 0.0035 & 0.0013 \\ 0 & 0 & 0 & 0 & -0.3122 & 0.1999 & 0.1999 & 0.1999 & 0.1999 \\ -0.0071 & -0.0071 & -0.0074 & -0.0074 & -0.0676 & -0.0004 & -0.0004 & -0.0004 & -0.0004 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$C = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & -0.0199 & 3.0796 & -9.8048 & 8.98\cdot 10^{-5} \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -92.6 & 92.6 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$D_u = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.3122 & 0.1999 & 0.1999 & 0.1999 & 0.1999 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$