

Redundancy Relations and Robust Failure Detection[†]

Edward Y. Chow^{††}

Xi-Cheng Lou*

George C. Verghese**

Alan S. Willsky*

Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Abstract

All failure detection methods are based on the use of redundancy, that is on (possible dynamic) relations among the measured variables. Consequently the robustness of the failure detection process depends to a great degree on the reliability of the redundancy relations given the inevitable presence of model uncertainties. In this paper we address the problem of determining redundancy relations which are optimally robust in a sense which includes the major issues of importance in practical failure detection and which provides us with a significant amount of intuition concerning the geometry of robust failure detection.

[†] This research was supported in part by The Office of Naval Research under Grant N00014-77-C-0224 and by NASA Ames and NASA Langley Research Centers under Grant NGL-22-009-124.

^{††} Presently with Schlumberger-Doll Research, P.O. Box, 307, Ridgefield, Conn. 06877.

* Also affiliated with the M.I.T. Laboratory for Information and Decision Systems.

** Also affiliated with the M.I.T. Electric Power Systems Engineering Laboratory.

1. Introduction

In this paper we consider the issue of robust failure detection. In one way or another all failure detection methods generate signals which tend to highlight the presence of particular failures if they have actually occurred. However, if any model uncertainties have effects on the observables which are at all like those of one or more of the failure modes, these will also be accentuated. Consequently the problem of robust failure detection is concerned with generating signals which are maximally sensitive to some effects (failures) and minimally sensitive to others (model errors).

The initial impetus for our approach to this problem came from the work reported in [5, 13] which document the first and to date by far most successful application and flight testing of a failure detection algorithm based on advanced methods which use analytic redundancy. The singular feature of that project was that the dynamics of the aircraft were decomposed in order to analyze the relative reliability of each individual source of potentially useful failure detection information.

In [2] we presented the results of our initial attempt to extract the essence of the method used in [5, 13] in order to develop a general approach to robust failure detection. As discussed in that reference and in others (such as [3, 7-9]), all failure detection systems are based on exploiting analytical redundancy relations or (generalized) parity checks. These are simply functions of the temporal histories of the measured quantities which have the property that they are small (ideally zero) when the system is operating normally. In [2] we present one criterion for measuring the reliability of a particular redundancy relation and use this to pose an optimization problem to determine the most reliable relation. In [3, 19] we present another method which has some computational advantages not found

in the approach described in [2].

In this paper we describe the major results of [2, 3, 19]. In the next section we review the notion of analytic redundancy for perfectly known models and provide a geometric interpretation which forms the starting point for our investigation of robust failure detection. Section 3 addresses the problem of robustness using our geometric ideas, and in that section we pose and solve a first version of the optimum robust redundancy problem. In Section 4 we discuss extensions to include three important issues not included in Section 3: scaling, noise, and the detection/robustness tradeoff.

2. Redundancy Relations

Consider the noise-free discrete-time model

$$x(k+1) = Ax(k) + Bu(k) \quad (2.1)$$

$$y(k) = Cx(k) \quad (2.2)$$

where x is n -dimensional, u is m -dimensional, y is r -dimensional, and A , B , and C are perfectly known. A redundancy relation for this model is some linear combination of present and lagged values of u and y which should be identically zero if no changes (i.e. failures) occur in (2.1), (2.2). As discussed in [2, 3, 19], redundancy relations can be specified mathematically in the following way. The subspace of $(p+1)r$ -dimensional vectors given by

$$G \triangleq \left\{ \omega \mid \omega' \begin{bmatrix} C \\ CA \\ \vdots \\ CA^p \end{bmatrix} = 0 \right\} \quad (2.3)$$

is called the space of parity or redundancy relations of order p . The reason for this terminology is the following. Suppose that $\omega \in G$. Then (2.1) - (2.3) imply that if we partition ω into $(p+1)$ subvectors of dimension r

$$\omega' = [\omega'_0, \dots, \omega'_p] \quad (2.4)$$

then at any time k

$$r(k) = \sum_{i=0}^p \omega'_i [y(k-p+i) - \sum_{j=0}^{i-1} CA^{i-j-1} Bu(k-p+j)] = 0 \quad (2.5)$$

The quantity $r(k)$ is called a parity check. A simpler form for (2.5)

(which we will use later) can be written in the case when $u = 0$ (or, equivalently, if the effect of the inputs are subtracted from the observations before computing the parity check). In this case

$$r(k) = \omega' \begin{bmatrix} y(k-p) \\ y(k-p+1) \\ \vdots \\ y(k) \end{bmatrix} \quad (2.6)$$

To continue our development, let us assume that

$$\omega_p \neq 0 \quad (2.7)$$

Let us denote the components of ω_i as

$$\omega_i' = [\omega_{i1}, \dots, \omega_{ir}] \quad (2.8)$$

Since at least one element of ω_p is nonzero, we can normalize ω so this component has unity value. In order to illustrate several points, let us assume that the first component, $\omega_{p1} = 1$. In this case (2.5) can be re-written as

$$\begin{aligned} y_1(k) = & - \sum_{i=0}^{p-1} \omega_{i1} y_1(k-p+i) - \sum_{i=0}^p \sum_{s=2}^r \omega_{is} y_s(k-p+i) \\ & + \sum_{i=0}^p \sum_{j=0}^{i-1} \omega'_i CA^{i-j-1} Bu(k-p+j) \end{aligned} = 0 \quad (2.9)$$

There are two very important interpretations of (2.9). The most obvious is that the right-hand side of this equation represents a synthetic measurement which can be directly compared to $y_1(k)$ in a simple comparison test. The second interpretation of (2.9) is as a reduced-order dynamic model. Specifically this equation is nothing but an autoregressive-moving average (ARMA) model for $y_1(k)$. (From the point of view of the evolution of y_1 according to (2.9), y_2, \dots, y_r and the components of u are all regarded as inputs). This second interpretation, allows us to make contact with the numerous existing failure detection methods. Typically such methods are based on a noisy version of the model (2.1), (2.2) representing normal system behavior together with a set of deviations from this model

representing the several failure modes. Rather than applying such methods to a single, all-encompassing model as in (2.1), (2.2), one could alternatively apply the same techniques to individual models as in (2.9) (or a combination of several of these), thereby isolating individual (or specific groups of) parity relations. For example, this is precisely what was done in [5, 13]. The advantage of such an approach is that it allows one to separate the information provided by redundancy relations of differing levels of reliability, something that is not easily done when one starts with the overall model (2.1), (2.2) which combines all redundancy relations.

In the next two sections we address the main problem of this paper, which is the determination of optimally robust redundancy relations. The key to this approach is the observation that G in (2.3) is the orthogonal complement of the range Z of the matrix

$$\begin{bmatrix} C \\ \vdots \\ CA^p \end{bmatrix} \tag{2.10}$$

Thus (assuming $u = 0$ or that the effect of u is subtracted from the observations) a complete set of independent parity relations of order p is given by the orthogonal projection of the window of observations $y(k)$, $y(k-1), \dots, y(k-p)$ onto G .

3. An Angular Measure of Robustness

Consider a model containing imperfectly known parameters η , process noise w and measurement noise v :

$$x(k+1) = A(\eta)x(k) + B(\eta)u(k) + w(k) \quad (3.1)$$

$$y(k) = C(\eta)x(k) + v(k) \quad (3.2)$$

where η is a vector of unknown parameters and where the matrices A , B , C and the covariances of w and v are functions of η . Let K denote the set of possible values which η can take on. In their work [2] Chow and Willsky used the following line of reasoning. If the parameters of the system were known perfectly and if there were no process or measurement noises, then according to (2.5) we could find a vector $\omega' = [\omega'_0, \dots, \omega'_p]$ and a vector $\mu = [\mu_0, \dots, \mu_{p-1}]$ with

$$\mu'_i = \sum_{j=i+1}^p \omega'_j CA^{j-i-1}B \quad (3.3)$$

so that

$$r(k) = \sum_{i=0}^p \omega'_i y(k-p+i) - \sum_{i=0}^{p-1} \mu'_i u(k-p+1) = 0 \quad (3.4)$$

In the uncertain case, what would seem to make sense is to minimize some measure of the size of $r(k)$. For example one could consider choosing ω and μ that solve the minimax problem

$$\min_{\omega, \mu} \max_{\eta \in K} E_{x_0(\eta), u_0} [r(k)]^2 \quad (3.5)$$

$\|\omega\| = 1$

Here the expectation is taken for each value of η and assuming that the system is at particular operating point, i.e. that $u(k) \equiv u_0$ and that $x_0(\eta)$ is the corresponding set point value of the state. This criterion has the

interpretation of finding the approximate parity relation which, at the specified operating point, produces the residual with the smallest worst-case mean-square value when no failure has occurred.

Let us make several comments concerning the procedure just described. In the first place the optimization problem (19) is a complex nonlinear programming problem. Furthermore, the method does not easily give a sequence of parity relations ordered by their robustness. Finally the optimum parity relation clearly depends upon the operating point as specified by u_0 and $x_0(\eta)$. In some problems this may be desirable as it does allow one to adapt the failure detection algorithm to changing conditions, but in others it might be acceptable or preferable to have a single set of parity relations for all operating conditions. The approach developed in this paper produces such a set and results in a far simpler computational procedure.

To begin, let us focus on (3.1), (3.2) with $u = w = v = 0$. Referring to the previous discussion, we note that it is in general impossible to find parity checks which are perfect for all possible values of η . That is, in general we cannot find a subspace G which is orthogonal to

$$Z(\eta) = \text{Range} \begin{bmatrix} C(\eta) \\ C(\eta)A(\eta) \\ \vdots \\ C(\eta)A(\eta)^P \end{bmatrix} \quad (3.6)$$

for all η .

What would seem to make sense in this case is to choose a subspace G which is "as orthogonal as possible" to all possible $Z(\eta)$. Several possible ways in which this can be done are described in detail in [3]. In this paper we focus on the one approach which leads to the most complete picture of robust redundancy and which is computationally the simplest. To do this, however, we must make the assumption that K , the set of possible values of

η is finite. Typically what this would involve is choosing representative points out of the actual, continuous range of parameter values. Here "representative" means spanning the range of possible values and having density variations reflecting any desired weightings on the likelihood or importance of particular sets of parameter values. However this is accomplished, we will assume for the remainder of this paper that η takes on a discrete set of values $\eta=1, \dots, L$, and will use the notation A_i for $A(\eta=i)$, Z_i for $Z(\eta=i)$, etc.

To obtain a simple computational procedure for determining robust redundancy relations we first compute an average observation subspace Z_0 which is as close as possible to all of the Z_i , and we then choose G to be the orthogonal complement of Z_0 . To be more precise, note first that the Z_i are subspaces of possibly differing dimensions ($\dim Z_i = v_i$) embedded in a space of dimension $N = (p+1)r$. We will find it convenient to use the same symbols Z_1, \dots, Z_L to denote matrices of sizes $N \times v_i$, $i=1, \dots, L$, whose columns form orthonormal bases for the corresponding subspaces. Letting $M = v_1 + \dots + v_L$, we define the $N \times M$ matrix

$$Z = [Z_1 \quad \vdots \quad \dots \quad \vdots \quad Z_L] \tag{3.7}$$

Thus the columns of Z span the possible directions in which observation histories may lie under normal conditions.

We now suppose that we wish to determine the s best parity checks (so that $\dim G=s$). Thus we wish to determine a subspace Z_0 of dimension $N-s$. The optimum choice for this subspace is taken to be the span of the (not necessarily orthogonal) columns of the matrix Z_0 which minimizes

$$\|Z - Z_0\|_F^2 \tag{3.8}$$

subject to the constraint that $\text{rank } Z_0 = N-s$. Here $\|\cdot\|_F$ denotes the Frobenius norm:

$$\|D\|_F^2 = \sum_j \sum_i |d_{ij}|^2 \quad (3.9)$$

There are several important reasons for choosing this criterion, one being that it does produce a space which is as close as possible to a specified set of directions. A second is that the resulting optimization problem is easy to solve. In particular, let the singular value decomposition of Z [14, 15] be given by

$$Z = U \Sigma V \quad (3.10)$$

where U and V are orthogonal matrices, and

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdot & & \\ & \cdot & & & \\ & & \cdot & & \\ 0 & & & \sigma_n & \cdot \\ & & & & & & 0 \end{bmatrix} \quad (3.11)$$

Here $\sigma_1 \leq \sigma_2 \leq \dots \leq \sigma_N$ are the singular values of Z ordered by magnitude. Note we have assumed $N \leq M$. If this is not the case we can make it so without changing the optimum choice of Z_0 by padding Z with additional columns of zeros. It is readily shown [17, 18] that the matrix Z_0 minimizing (3.8) is given by

$$Z_0 = U \begin{bmatrix} 0 & & & & 0 & & & & \\ & \cdot & & & & & & & \\ & & \cdot & & & & & & \\ & & & \cdot & & & & & \\ & & & & \sigma_{s+1} & & & & \\ & & & & & \cdot & & & \\ & & & & & & \cdot & & \\ & & & & & & & \sigma_N & \\ 0 & & & & & & & & & & \cdot & \\ & & & & & & & & & & & 0 \end{bmatrix} V \quad (3.12)$$

Moreover, since the columns of U are orthonormal, we immediately see that the orthogonal complement of the range of Z_0 is given by the first s left singular vectors of Z_0 , i.e. the first s columns of U . Consequently

$$G = [u_1 \vdots \dots \vdots u_s] \quad (3.13)$$

and u_1, \dots, u_s are the optimum redundancy relations.

There is an alternative interpretation of this choice of G which

provides some very useful insight. Specifically, recall that what we wish to do is to find a G whose columns are as orthogonal as possible to the columns of the Z_i ; that is, we would like to choose G to make each of the matrices $Z_i'G$ as close to zero as possible. In fact, as shown in [3], the choice of G given in (3.13) minimizes

$$J(s) = \sum_{i=1}^L \|Z_i'G\|_F^2 \quad (3.14)$$

yielding the minimum value

$$J(s) = \sum_{i=1}^S \sigma_i^2 \quad (3.15)$$

There are two important points to observe about the result (3.14), (3.15). The first is that we can now see a straightforward way in which to include unequal weightings on each of the terms in (3.14). Specifically, if the w_i are positive numbers, then

$$\sum_{i=1}^L w_i \|Z_i'G\|_F^2 = \sum_{i=1}^L \|\sqrt{w_i} Z_i'G\|_F^2 \quad (3.16)$$

so that minimizing this quantity is accomplished using the same procedure described previously but with Z_i replaced by $\sqrt{w_i} Z_i$. As a second point note that the optimum value (3.17) provides us with an interpretation of the singular values as measures of robustness and with an ordered sequence of parity relations from most to least robust.

4. Several Important Extensions

In this section we address several of the drawbacks and limitations of the result of the preceding section and obtain modifications to this result which overcome them at no fundamental increase in complexity.

4.1 Scaling

A critical problem with the method used in the preceding section is that all vectors in the observation spaces Z_i are treated as being equally likely to occur. If there are differences in scale among the system variables this may lead to poor solutions for the optimum parity relations. To overcome this drawback we proceed as follows. Suppose that we are given a scaling matrix P so that with the change of basis

$$\xi = Px \tag{4.1}$$

one obtains a variable ξ which is equally likely to lie in any direction. For example if covariance analysis has been performed on x and its covariance is Q , then P can be chosen to satisfy

$$Q = P^{-1}(P')^{-1} \tag{4.2}$$

and the resulting covariance of ξ is the identity.

As a next step, recall that what we would ideally like to do is to choose a matrix G so that

$$G' \begin{bmatrix} C_i \\ C_i A_i \\ \vdots \\ C_i A_i P \end{bmatrix} x = G' \begin{bmatrix} C_i P^{-1} \\ C_i A_i P^{-1} \\ \vdots \\ C_i A_i P^{-1} \end{bmatrix} \xi \triangleq G' \bar{C}_i \xi \tag{4.3}$$

is as small as possible. In the preceding section we considered all directions in $Z_i = \text{Range}(\bar{C}_i)$ to be on equal footing and arrived at the criterion (4.4)

Since all directions for ξ are on equal footing, we are led naturally to the following criterion which takes scaling into account

$$J(s) = \sum_{i=1}^L \|\bar{C}_i' G\|_F^2 \quad (4.4)$$

Using the result [17] cited in the previous section we see that to find the $N \times s$ matrix G (with orthonormal columns) which minimizes $J(s)$ we must perform a singular value decomposition of the matrix

$$\bar{C} = [\bar{C}_1 : \bar{C}_2 : \dots : \bar{C}_L] = U \Sigma V \quad (4.5)$$

where $\sigma_1^2 \leq \sigma_2^2 \leq \dots \leq \sigma_N^2$ and $U = [u_1 : u_2 : \dots : u_N]$. Then u_1 is the best parity relation with σ_1^2 as its measure of robustness, u_2 is the next best, etc., and $J^*(s)$ is given by (3.15). Finally, in anticipation of the next subsection, suppose that we use the stochastic interpretation of ξ , i.e. that

$$E[\xi\xi'] = I \quad (4.6)$$

In this case if we define the parity check vector

$$\mu_i = G' \bar{C}_i \xi \quad (4.7)$$

then

$$E[\|\mu_i\|^2] = \|C_i' G\|_F^2 \quad (4.8)$$

4.2 Observation and Process Noise

In addition to choosing parity relations which are maximally insensitive to model uncertainties it is also important to choose relations which suppress noise. Consider then the model

$$x(k+1) = A_i x(k) + D_i w(k) \quad (4.9)$$

$$y(k) = C_i x(k) + v(k) \quad (4.10)$$

where w and v are independent, zero-mean white noise processes with covariances

Q and R, respectively.

Let

$$\mu = G' \begin{bmatrix} y(k) \\ \vdots \\ y(k+p) \end{bmatrix} \quad (4.10)$$

Then using the interpretation provided in (4.7), we obtain the following natural generalization of the criterion (4.4):

$$J(s) = \sum_{i=1}^L E_i [\|\mu\|^2] \quad (4.11)$$

where E_i denotes expectation assuming that the i th model is correct. Assuming that $\xi(k) = Px(k)$ has the identity as its covariance, using the whiteness of w and v , and performing some algebraic manipulations we obtain [3]

$$J(s) = \sum_{i=1}^L \|C_i' G\|_F^2 + \|S' G\|_F^2 \quad (4.12)$$

where S is defined by the following:

$$\bar{D}_i = \begin{bmatrix} 0 & 0 & \dots & 0 \\ C_i D_i & 0 & & \vdots \\ C_i A_i D_i & C_i D_i & & \vdots \\ \vdots & \vdots & & 0 \\ C_i A_i^{p-1} D_i & C_i A_i^{p-2} D_i & \dots & C_i D_i \end{bmatrix} \quad (4.13)$$

$$\bar{Q} = \text{diag} (Q, \dots, Q) \text{ (p times)}$$

$$\bar{R} = \text{diag} (R, \dots, R) \text{ ((p+1) times)} \quad (4.14)$$

$$N = \sum_{i=1}^L \bar{D}_i \bar{Q} \bar{D}_i' + \bar{L} \bar{R} = SS' \quad (4.15)$$

From (4.12) we see that the effect of the noise is to specify another set of directions, namely the columns of S , to which we would like to make the columns of G as close to orthogonal as possible. From this it is evident

that the optimum choice of G is computed by performing a singular value decomposition on the matrix

$$[\bar{C}_1 \vdots \dots \vdots \bar{C}_L \vdots S] = U \Sigma V \quad (4.16)$$

As before (4.16) provides a complete set of parity relations ordered in terms of their degrees of insensitivity to model errors and noise.

4.3 Detection Versus Robustness

The methods described to this point involve measuring the quality of redundancy relations in terms of how small the resulting parity checks are under normal operating conditions. However, in some cases one might prefer to use an alternative viewpoint. In particular there may be parity checks which are not optimally robust in the senses we have discussed but are still of significant value because they are extremely sensitive to particular failure modes. In this subsection we consider a criterion which takes such a possibility into account. For simplicity we focus on the noise-free case. The extension to include noise as in the previous subsection is straightforward.

The specific problem we consider is the choice of parity checks for the robust detection of a particular failure mode. We assume that the unfailed model of the system is

$$x(k+1) = A_u(\eta)x(k) \quad (4.17)$$

$$y(k) = C_u(\eta)x(k) \quad (4.18)$$

while if the failure has occurred the model is

$$x(k+1) = A_f(\eta)x(k) \quad (4.19)$$

$$y(k) = C_f(\eta)x(k) \quad (4.20)$$

In this case one would like to choose G to be "as orthogonal as possible" to

$Z_u(\eta)$ and "as parallel as possible" to $Z_f(\eta)$.

Assume again that η takes on one of a finite set of possible values, and let \bar{C}_{ui} and \bar{C}_{fi} denote the counterparts of \bar{C}_i in (4.3) for the unfailed and failed models, respectively. A natural criterion which reflects our objective is

$$J(s) = \min_{G'G=I} \sum_{i=1}^L \{ \|\bar{C}'_{ui} G\|_F^2 - \|\bar{C}'_{fi} G\|_F^2 \} \quad (4.21)$$

If we define the matrix

$$H = \underbrace{[\bar{C}_{u1} \quad \bar{C}_{u2} \quad \dots \quad \bar{C}_{uL}]_{M_1 \text{ columns}}}_{M_1 \text{ columns}} \quad \underbrace{[\bar{C}_{f1} \quad \bar{C}_{f2} \quad \dots \quad \bar{C}_{fL}]_{M_2 \text{ columns}}}_{M_2 \text{ columns}} \quad (4.22)$$

$$J(s) = \min_{G'G=I} \text{tr}\{G'HS'HG\} \quad (4.23)$$

where

$$S = \begin{bmatrix} -I & \vdots & 0 \\ \dots & \ddots & \dots \\ 0 & \vdots & I \end{bmatrix} \begin{matrix} M_1 \\ M_2 \end{matrix} \quad (4.24)$$

It is straightforward (see [3]) to show that a minor modification of the result in [17] leads to the following solution. We perform an eigenvector-eigenvalue analysis on the matrix

$$HS'H = U \Lambda U' \quad (4.25)$$

where $U'U = I$ and

$$\Lambda = \text{diag} (\lambda_1, \dots, \lambda_N) \quad (4.26)$$

with $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_N$ and $U = [u_1 \quad \dots \quad u_N]$. Then the optimum choice of G is

$$G = [u_1 \quad \dots \quad u_s] \quad (4.27)$$

and the corresponding value of (4.23) is

$$J^*(s) = \sum_{i=1}^s \lambda_i \quad (4.28)$$

Let us make two comments about this solution. The first is that as many as M_2 of the λ_i can be negative. In fact the parity check based on u_i is likely to have larger values under failed rather than unfailed conditions if and only if $\lambda_i < 0$. Thus we immediately see that the maximum number of useful parity relations for detecting this particular failure mode equals the number of negative eigenvalues of HS^H . As a second comment, let us contrast the procedure we use here with a singular value decomposition, which corresponds essentially to performing an eigenvector-eigenvalue analysis of HH' . First, assume that the first K of the λ_i are negative. Then, define

$$\begin{aligned} \sigma_1^2 &= -\lambda_1, \sigma_2^2 = -\lambda_2, \dots, \sigma_K^2 = -\lambda_K, \\ \sigma_{K+1}^2 &= \lambda_{K+1}, \dots, \sigma_N^2 = \lambda_N \end{aligned} \quad (4.29)$$

From (4.25) we have that

$$HS^H = U\Sigma S^H U'$$

where

$$\Sigma = \text{diag}(\sigma_1, \dots, \sigma_N) \quad (4.31)$$

Assuming that Σ is nonsingular, define

$$V = \Sigma^{-1} U^H \quad (4.32)$$

Then (4.31), (4.32) imply that V is S-orthogonal

$$V S V^H = S \quad (4.33)$$

and that H has what we call as S-singular value decomposition

$$H = U \Sigma V \quad (4.34)$$

REFERENCES

1. A.S. Willsky, "A Survey of Design Methods for Failure Detection," Automatica, Vol. 12, 1976, pp. 601-611.
2. E.Y. Chow and A.S. Willsky, "Analytical Redundancy and the Design of Robust Failure Detection Systems," to appear in the IEEE Trans. on Aut. Control.
3. X.-C. Lou, "A System Failure Detection Method: The Failure Projection Method," S.M. Thesis M.I.T. Dept. of Elec. Eng. and Comp. Sci.; also Rept. LIDS-TH-1203, M.I.T. Lab. for Inf. and Dec. Sys., June 1982.
4. A.S. Willsky, "Failure Detection in Dynamic Systems," AGARD Lecture Series No. 109 on Fault Tolerance Design and Redundancy Management Techniques," Athens, Rome, and London, Oct. 1980.
5. J.C. Deckert, M.N. Desai, J.J. Deyst, and A.S. Willsky, "F8-DFBW Sensor Failure Detection and Identification Using Analytic Redundancy," IEEE Trans. Aut. Control, Vol. AC-22, No. 5, Oct. 1977, pp. 795-803.
6. G.C. Leininger, "Sensor Failure Detection with Model Mismatch," 1981 Joint Automatic Control Conf., Charlottesville, Virginia, June 1981.
7. S.R. Hall, "Parity Vector Compensation for FDI," S.M. Thesis, Dept. of Aero. and Astro., M.I.T., Feb. 1981.
8. J.E. Potter and M.C. Suman, "Thresholdless Redundancy Management with Arrays of Skewed Instruments," in AGARDOGRAPH-224 on Integrity in Electronic Flight Control Systems, 1977, pp. 15-1 to 15-25.
9. M. Desai and A. Ray, "A Failure Detection and Isolation Methodology," 20th IEEE Conf. on Decision and Control, San Diego, California, Dec. 1981.
10. F.C. Schweppe, Uncertain Dynamic Systems, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1973.
11. J.R. Winkelman, "Multi-time-scale Analysis of a Power System," Automatica, Vol. 16, 1980, pp. 35-43.
12. A.S. Willsky, E.Y. Chow, S.B. Gershwin, C.S. Greene, P.K. Houpt, and A.L. Kurkjian, "Dynamic Model-Based Techniques for the Detection of Incidents on Freeways," IEEE Trans. Aut. Control, Vol. AC-25, No. 3, June 1980, pp. 347-360.
13. J.C. Deckert, "Flight Test Results for the F-8 Digital Fly-by-Wire Aircraft Control Sensor Analytic Redundancy Management Technique," Proc. AIAA Guidance and Control Conference, Albuquerque, New Mexico, Aug. 1981, pp. 491-499.
14. V.C. Klema and A.J. Laub, "The Singular Value Decomposition: Its Computation and Some Applications," IEEE Trans. Aut. Control, Vol. AC-25, No. 2, April 1980, pp. 164-176.
15. B. Noble, Applied Linear Algebra, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1969.

16. A. Bjorck and G.H. Golub, "Numerical Methods for Computing Angles Between Linear Subspaces," Math. Comput., Vol. 27, 1973, pp. 579-594.
17. C. Eckart and G. Young, "The Approximation of One Matrix by Another of Lower Rank," Psychometrika, Vol. 1, 1936, pp. 211-218.
18. D.W. Tufts, R. Kumaresan, and I. Kirsteins, "Data Adaptive Signal Estimation by Singular Value Decomposition of a Data Matrix," Proc. IEEE, Vol. 70, No. 6, June 1983, pp. 684-685.
19. X.-C. Lou, A.S. Willsky, and G.C. Verghese, "Failure Detection with Uncertain Models, " Proc. 1983 Amer. Control Conf., San Francisco, June 1983.