A New Structural Framework for Parity Equation-based Failure Detection and Isolation*

JANOS GERTLER† and DAVID SINGER‡

Key Words—Failure detection; diagnosis; isolability; analytical redundancy; parity equations.

Abstract—The paper describes a new framework for developing parity equations that prevent incorrect isolation decisions under marginal size failures in a decision process that tests each residual independently. Test thresholds that take the noise conditions into account are set high to reduce the occurrence of false alarms while maintaining the algorithm's ability to detect and isolate larger failures. The method is applicable to additive failures on the measured input and output variables and to additive plant disturbances. A transformation algorithm provides a multitude of models that satisfy the isolability requirements. A search procedure utilizing this model redundancy integrates model robustness considerations into the design.

Introduction

The early and reliable indication of equipment failures is of great significance from the point of view of operational safety. The task of failure detection and isolation (FDI) systems includes the detection of the presence of failure and the isolation of the component responsible for the irregularity. The FDI activity is to cover both the basic equipment and its control instrumentation (sensors and actuators). The main applications of FDI can be found in the process industries (chemical plants, oil refineries, power stations) and in the aerospace industry (high-performance airplanes, space vehicles and stations); and most recently also in some mass-produced consumer equipment (automobiles, appliances).

With the proliferation of on-line computers, a wide variety of FDI approaches have been developed. These include simple limit-checking, redundant sensors, special sensors, frequency analysis and artificial intelligence techniques. The most important class of FDI methods is based on the concept of analytical redundancy. Analytical redundancy techniques make use of an explicit model of the plant; sensory observations are confronted with the prior information embodied in the model. Reviews of model-based FDI techniques can be found in survey papers by Willsky (1976), Isermann (1984) and Gertler (1986a).

Model-based FDI techniques usually consist of three stages (Chow and Willsky, 1984). In the first stage, residuals are generated, using the model and sensory observations. Residuals are variables that are zero under ideal circumstances; they become nonzero as a result of failures, noise and modeling errors. To account for the presence of noise, the residuals are subjected to statistical testing, yielding failure signatures. Finally, these signatures are logically analyzed to arrive at a failure inference.

Several principles may be used to generate residuals. These include Kalman filtering, observers and the straightforward use of an input-output model. In this paper, only the input-output model approach will be discussed.

In the chemical engineering literature, usually material and energy balance equations serve as residual generating input-output models. These express the laws of conservation and are generally static. Vaclavek (1974) proposed parallel statistical tests on the short-term averages of balance equation residuals. Almasy and Szanto (1975) detected bias faults on the basis of the size of elements in a transformed residual vector. Mah et al. (1976) introduced a systematic sequential algorithm to isolate faults in networks based on balance equation residuals around groups of nodes. In the method of Romagnoli and Stepanopoulos (1981), a single statistic derived from balance equation residuals was subjected to statistical testing in a sequential procedure. Stanley and Mah (1977) and Rooney et al. (1978) applied Kalman filtering to balance equation residuals. Ben Haim (1980) used redundant balance equations for residual generation and formulated the structural requirement, referred to later in this paper as the zero-threshold isolability.

In the aerospace and related control engineering literature, the residual generating input-output models are usually dynamic and are referred to as parity equations. Parity equations have been used in FDI systems for inertial navigation (Satin and Gates, 1978) where relationships between readings in gyroscope/accelerometer assemblies provide analytical redundancy (static equations). A combination of physical and analytical redundancy was applied in the FDI system of the NASA F-8 digital fly-by-wire aircraft (Dockert et al., 1977), the dynamic system equations serving as the parity relationships. A geometric theory of (static) parity equations was presented by Potter and Suoman (1977).

Owing to inevitable inaccuracies in the model, robustness relative to modeling errors is a critical issue of FDI design. Based on the wealth of experience gained with the F-8 project, Chow and Willsky (1984) proposed a systematic mechanism to generate parity equations and formulated a minimax framework for robust design. Lou et al. (1986) recently developed a design methodology that provides the most robust parity equations, given a finite set of uncertain plant models.

In the method to be described here, the residuals are obtained from the dynamic input-output equations of the system. These residuals are subjected to statistical testing individually, in parallel. The outcomes of the tests are

* Received 24 October 1985; revised 4 March 1987; revised 6 January 1988; revised 19 July 1988; revised 22 March 1989; received in final form 23 May 1989. The original version of this paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor B. Freidland under the direction of Editor H. Austin Spang III.

† Corresponding author. Department of Electrical and Computer Engineering, George Mason University, Fairfax, VA 22030, U.S.A.

‡ Department of Mathematics and Statistics, Case Western Reserve University, Cleveland, OH 44106, U.S.A.
combined into a Boolean vector that constitutes the failure signature.

The types of failure handled by the proposed method include single additive failures of the measured input and output variables and single additive plant disturbances. For the latter, knowledge of the system's disturbance model is also required, in addition to the input-output model. The technique may be extended to multiplicative (model parameter) type failures.

Attention in this paper is focused on the isolability properties of the failure detection and isolation algorithm. Residuals that are orthogonal to certain additive failures are considered. Such orthogonality places structural constraints on the parity equations. Building from these equations, model structures that show advantageous isolability properties are defined. One such structure guarantees the isolability of the single additive failures if the test thresholds are set to zero (zero-threshold isolability). Another structure prevents incorrect failure inferences that could result from marginal size failures with the thresholds set high (high-threshold isolability).

A transformation technique is introduced that can be used to generate the models required for isolability. Special properties (zeros and dependences) of the original system model place certain restrictions on the transformation procedure. In spite of such restrictions, usually there are a combinatorially large number of model structures satisfying the isolability conditions (model redundancy). This allows for the incorporation of robustness considerations with respect to modeling errors. A design procedure along these lines is briefly described.

The idea of model transformation with isolable structures in mind appeared in a work by Ben-Haim (1983); however, he was only concerned with zero-threshold isolability. The orthogonal residuals required for isolability could also be obtained by the residual generation technique of Chow and Willsky (1984), by imposing structural constraints, but the simplicity of the input-output model would be lost. The robustness considerations offered in this paper are, in a sense, elaborations on some ideas also appearing in the latter reference. The approach proposed here is aimed at a problem similar to the one solved by Lou et al. (1986) but, unlike their approach, with the focus on isolability.

Preliminary results of the work discussed here were reported in several conference papers (Gertler and Singer, 1985; Gertler et al., 1985; Gertler, 1988b). Some ideas first appeared in research reports (Shutty, 1985; Sundar, 1985).

System description

We will concern ourselves with multiple input, multiple output discrete linear dynamic systems. This assumption implies that continuous-time systems are discretized and nonlinear systems are linearized. Static systems are included as a special case.

Consider an input vector \( u(t) = [u_1(t), \ldots, u_n(t)]^T \) and an output vector \( y(t) = [y_1(t), \ldots, y_m(t)]^T \). The ideal ARMA system is described by its input-output relationship

\[
B(z)u(t) = A(z)y(t)
\]

where

\[
B(z) = B_0 + B_1 z^{-1} + \cdots + B_n z^{-n}
\]

\[
A(z) = A_0 + A_1 z^{-1} + \cdots + A_n z^{-n}
\]

are known polynomial system transfer matrices in the shift-operator \( z^{-1} \) and \( n \) is the system order. Matrices \( A(z) \) and \( B(z) \) are row-wise relative prime (contain no common factor) and \( A(z) \) is diagonal. This is the simplest input-output description of a system that contains only controllable and observable modes.

To obtain parity equation formulation, introduce

\[
x(t) = [u(t) \quad y(t)]^T, \quad F(z) = [B(z) \quad -A(z)].
\]

With these, the system equation (1) becomes

\[
F(z)x(t) = 0.
\]

The real system differs from the ideal one, first of all, in that the outputs depend also on disturbance inputs. The disturbance inputs will be divided into two groups: \( w(t) \), the zero mean random disturbance inputs representing process noise, and \( v(t) = [v_1(t), \ldots, v_n(t)]^T \), the semideterministic disturbance inputs (e.g., steps occurring at random times) representing additive faults in the plant (e.g., leaks, certain actuator failures). With these, system equation (1) becomes

\[
B(z)w(t) + G(z)v(t) + H(z)w(t) = A(z)y(t)
\]

where \( G(z) \) and \( H(z) \) are the known disturbance transfer matrices of the system.

Further, the system variables \( u(t) \) and \( y(t) \) cannot be observed exactly due to observation noise \( \delta u(t), \delta y(t) \) and to any measurement bias \( \Delta u(t), \Delta y(t) \). Again, the noise is assumed to be random with zero mean and the bias is semideterministic. Thus, the observed input-output vector \( \tilde{x}(t) \) is

\[
\tilde{x}(t) = x(t) + \delta x(t) + \Delta x(t)
\]

where \( \Delta x(t) \) is the combined measurement bias and \( \delta x(t) \) is the combined measurement noise.

Finally, our perception of the system, embodied in the estimated system model \( \tilde{F}(z) \), is usually different from the true system matrix \( F(z) \):

\[
\tilde{F}(z) = F(z) + \Delta F(z).
\]

The discrepancy \( \Delta F(z) \) may be due to modeling errors or to changes in the plant proper (parametric plant faults like surface fouling, drop in pump power, etc.).

The parity equation (3) can, in reality, be applied only to the observations \( \tilde{x}(t) \) with the estimated system model \( \tilde{F}(z) \). This parity equation will generally yield a nonzero residual \( e(t) \):

\[
e(t) = \tilde{F}(z)\tilde{x}(t).
\]

Equation (7) is the external form of the parity equation showing how the residuals are computed. Expanding \( \tilde{F}(z) \) and \( \tilde{x}(t) \) yields the internal form of the residual equation:

\[
e(t) = -G(z)v(t) + H(z)w(t) + \Delta F(z)\tilde{x}(t)
\]

\[
+ \tilde{F}(z)\Delta x(t) + \tilde{F}(z)\delta x(t).
\]

This latter form reveals that the residuals depend on the noise \( w(t) \), \( \delta x(t) \), on the additive faults \( v(t) \), \( \delta x(t) \), and on the model discrepancy \( \Delta F(z) \).

Isolability conditions

The residuals obtained from the set of parity equations are subjected to statistical testing, in parallel, independent of each other. Such statistical testing implies the comparison of the momentary residuals to thresholds, previously established for each parity equation. Whether such a test fires or not depends on a number of factors. These include the model, the noise statistics and the selected "level of significance" (the probability of false alarms when there is no failure) which, together, determine the threshold. It also depends on the size of the failure, the momentary noise, and the "direction" of the residual relative to the failure.

To enhance the isolation of failures, it is desirable to work with residuals that are orthogonal to certain failures. A residual is orthogonal to a failure if the coefficient of that failure is identically zero in the parity equation. No matter how large a failure will never trigger the test on a residual that is orthogonal to it.

In the following, two isolability concepts will be introduced. Loosely speaking, a failure is considered isolable if it can be uniquely traced back from the results of the tests on the residuals. Both isolability concepts are based on orthogonal residuals. Clearly, orthogonality of residuals only depends on the position of zero coefficients, that is, on the structure of the parity equations. Some structural definitions will be given below.

The isolability concepts are rather straightforward for additive failures (measurement bias and plant disturbance).
The following treatment will be restricted to such failures, that is, $\Delta F = 0$ will be assumed. An extension to multiplicative failures has been outlined elsewhere (Gentler et al., 1985).

**Structural definitions.** The structure of matrix $F(z)$ is characterized by the Boolean incidence matrix

$$
\Phi = \text{inc} \{ F(z) \}
$$

(9)

where the elements $\phi_i$ of $\Phi$ are related to the elements $f_j$ of $F$ as $\phi_{ij} = 0$ if $f_j(0) = 0$ and $\phi_{ij} = 1$ if $f_j(z) = 0$. The $j$th column of the incidence matrix $\Phi$ will be denoted as $\phi_j$.

Similarly, the structure of matrix $G(z)$ is characterized by the incidence matrix $\Gamma$ with $\gamma_i$ being the $i$th column of $\Gamma$.

A system will be called row-canonical of order $i$ if each row of its incidence matrix contains $i$ zeros, each in different configuration. Similarly a system is column-canonical of order $i$ if each column of its incidence matrix contains $i$ zeros, each in different configuration.

A Boolean signature vector $e(t)$ is associated with the residual vector $e(t)$ according to the relationship $e(t) = 0$ if $e(t) \leq \eta_i$ and $e(t) = 1$ if $e(t) > \eta_i$, where $e(t)$ and $e(t)$ are elements of $e(t)$ and $e(t)$, respectively, and $\eta_i$ is the threshold specified for $e(t)$.

Zero-threshold (deterministic) isolaiblity. Assume that there is no plant and observation noise, that is $w = 0$, $\delta x = 0$. Then

$$
e(t) = G(0)v(t) + \hat{F}(0) \Delta x(t).
$$

(10)

In this case, the thresholds can be chosen as $\eta_i = 0$ for all $i$.

Each column of the incidence matrices $\Phi$ and $\Gamma$ is associated with an element of vectors $\Delta x(t)$ and $v(t)$, respectively. Any fault $\Delta x(t) \neq 0$ or $v(t) \neq 0$ results in $e(t) = 1$ if and only if $\phi_i = 1$ or $\gamma_i = 1$, respectively. That is, a single fault $\Delta x(t) \neq 0$ or $v(t) \neq 0$, respectively leads to

$$
e(t) = \phi_i \quad \text{or} \quad e(t) = \gamma_i.
$$

(11)

From this, it follows that:

(a) a single additive fault is undetectable if its column in the incidence matrix contains all zeros;

(b) in the zero noise situation, two single additive faults can be isolated from each other (that is, we can tell which one of the two faults occurs), if their two respective columns in the incidence matrix are different; and

(c) a system is zero-threshold isolable with respect to single additive faults if all columns of the incidence matrices are different and nonzero.

Note that the concept of zero-threshold isolability can be extended to multiple additive faults, making use of the fact that such faults result in signature patterns that are the Boolean sums of the respective columns of the incidence matrix.

**High-threshold (statistical) isolability.** Now consider the real system with plant and observation noise (but neglect the model discrepancy term). That is,

$$
e(t) = G(z)w(t) + \hat{F}(z) \Delta x(t) - H(z)w(t) + \hat{F}(z) \delta x(t).
$$

(12)

Assume that the noise $w(t)$ and $\delta x(t)$ are independent of $x(t)$, $v(t)$ and $\Delta x(t)$ and their distribution is completely known including their joint distribution for different time-shifts. Assume that all noises have zero mean.

To establish whether there is sufficient ground to assume that some biases are nonzero, the momentary residuals are subjected to statistical testing. Each component of the residual vector is tested separately, in parallel, using a simple threshold test. Since the residual vector is the output of a linear discrete moving average process with known parameters, the distribution of its components, with zero bias inputs, can be derived from the noise distributions. A confidence level can then be chosen and thresholds $\eta_i$ obtained for each $i$. Testing the actual residual values $e_i(t)$ against these thresholds yields the momentary Boolean signature vector $e(t)$.

Individual statistical tests may yield false inferences (false alarm or missed detection) with certain probability. False inferences cannot be avoided but their probability may be influenced by the selection of the test threshold. We conjecture here that the thresholds are set high so that false alarms are less likely than missed detections. This reflects the intuitive feeling that a missed detection is less of a problem than a false alarm, especially since a large fault will not really be missed even if the threshold is high. Similar considerations were adopted in some applications (in Satin and Gates, 1978), for example, the thresholds were set so that missed detections were twenty times as likely as false alarms.

A set of parallel tests is said to "fire completely" if a fault triggers the test on all the residuals that are not orthogonal to it. Under complete firing in response to a fault $\Delta x(t) \neq 0$ or $v(t) \neq 0$, the signature $e(t) = \phi_i$ or $e(t) = \gamma_i$, respectively, is obtained. Alternatively, a set of parallel tests "fires partially" if the test on some of the residuals not orthogonal to a particular fault is not triggered. The resulting Boolean signature $e(t)$ is a "degraded" version of $\phi_i$ or $\gamma_i$, respectively, in that some of the ones are replaced by zeros.

While large faults result in complete firing and small ones result in no firing at all, faults of intermediate size frequently cause partial firing, especially if the thresholds are set high. The danger of partial firing is the possibility of attributing the resulting degraded signature to another fault. The probability of this happening may be reduced considerably by some appropriate "high-threshold isolable" system structure, characterized by the following properties:

(a) A single additive fault $\Delta x_i \neq 0$ is high-threshold isolable (or statistically isolable) from another additive fault $\Delta x_j \neq 0$ if $\phi_i \neq \phi_j$ and $\phi_i$ cannot be obtained from $\phi_j$ by degradation. High-threshold isolability of $\Delta x_i \neq 0$ from $\Delta x_j \neq 0$, or of $\Delta x_i \neq 0$ from $\Delta x_j \neq 0$ requires similar properties of the respective columns of the incidence matrix. High-threshold isolability of two additive faults is not necessarily bidirectional.

(b) A system is high-threshold isolable with respect to single additive faults if it is zero-threshold isolable and no column of its incidence matrices can be obtained from any other column by degradation.

Isolability is posed above basically as a coding problem. What is required is that, between each pair of columns, $\phi_i$ and $\phi_j$, for example, there should be at least one position $i$ where $\phi_i = 1$ and $\phi_j = 0$ and another position $g$ where $\phi_i = 0$ and $\phi_j = 1$. While this can be achieved in a number of structures, a convenient sufficient condition is formulated below:

(c) Any column-canonical structure satisfies the conditions of high threshold isolability.

Obviously, all columns of a column-canonical system are different (since the configuration of zeros is different) and no column can be obtained from any other by degradation (since the number of zeros is the same).

**Example 1.** To illustrate the idea of partial firing and high-threshold isolability, consider a simple static system with no plant disturbance. Let the ideal input-output model be

$$
u(t) = y_2(t)
$$

(0.1) $u(t) = y_3(t)

The incidence matrix is

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

indicating that the system is isolable in the deterministic (zero-threshold) sense but not statistically (with high thresholds). The parity equations are

$$\begin{align*}
\Delta u(t) &+ \delta u(t) - \delta y_1(t) = e_3(t) \\
0.1 \Delta u(t) + 0.1 \delta u(t) - \delta y_2(t) &= e_3(t).
\end{align*}
$$
Assume that the noises $\delta u(t)$, $\delta y_1(t)$ and $\delta y_2(t)$ are normally distributed, independent of each other, have zero mean and unit variance. Select 90\% level for the statistical tests. The thresholds are obtained as

$$
\eta_1 = 1.65\sqrt{\text{Var} (e_1)} = 1.65\sqrt{(1 + 1)} = 2.35
$$

\( \eta_2 = 1.65\sqrt{\text{Var} (e_2)} = 1.65\sqrt{(0.01 + 1)} = 1.66 \).

Assume there is a bias $\Delta u(t) = 10$. Assume also that the other biases are zero and so are the momentary noise values. Thus,

$$
e_1(t) = \Delta u(t) = 10
e_2(t) = 0.1 \Delta u(t) = 1.
$$

That is, the test on $e_1(t)$ is triggered while that on $e_2(t)$ is not.

The $e(t) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ signature obtained is mistaken as resulting from $\Delta y(t) = 0$. Note that $\Delta u(t) > 16.6$ would fire both tests while $\Delta u(t) \leq 2.35$ would fire neither.

Although a high-threshold isolable structure does not eliminate false inferences from the individual statistical tests, it is a powerful tool in improving the robustness of the parallel fault isolation procedure. It reduces significantly the probability of one of the most frequent sources of misolation: the false inference from a parallel fault signature due to partial firing. Degraded signatures are simply neglected or are refined via statistical filtering of the residuals (Evans and Wilcox, 1970; Gertler and Singer, 1985).

The concept of statistical isolability can be extended to multiple bias faults, though the combinatorial dimensions may become prohibitive. Also, while the technique presented here provides for detecting and neglecting degraded signatures, it can be developed into one making the isolation of the right fault from (most) of such signatures possible as well ("error correcting" vs. "error detecting" codes).

**Model transformation**

Most system models in their original form do not satisfy the isolability conditions. It is possible, however, to generate new (secondary) equations by linear transformation applied to the equations of the original (primary) model. The secondary residuals generated this way will themselves be linear transformations of the primary ones. Transformed models may then be obtained as arbitrary combinations of primary and secondary equations. Thus, the structural requirements for isolability can generally be satisfied without altering the physical system.

The parity equations sought in the model generation/transformation procedure are orthogonal to a certain combination of failures; that is, they have zero coefficients at certain positions. Each of the primary equations, as seen in equations (1)–(3), is orthogonal to all but one output variable (more accurately, to the measurement bias associated with them). They may also be orthogonal to some of the input variables (measurement biases) and plant disturbances. The secondary equations are generated so that they be orthogonal to one or several input and/or plant disturbance faults. This is usually achieved at the expense of giving up orthogonality to a number of output faults.

The orthogonality requirements are formulated in terms of the internal parity equations [equation (8)]; the desired structures $\Phi^G$ and $\Phi'$ are specified. The transformation results in new matrices $F(x)$ and $G(x)$, of which only $F(x)$ is then used in the external parity equations (7).

To derive equations orthogonal to disturbance faults, $G(x)$ must be known. This may prove a significant constraint since the disturbance model $G(x)$ is usually much more difficult to obtain than the input–output model.

It is not always possible to perform the intended model transformation and attain the desired parity equations. Restrictions on the transformation procedure result from special properties of the $[G(x) \quad \hat{F}(x)]$ matrix of the original system model. Full zero columns, full column or row-dependence signify "major restrictions" of the system; these may require alteration of the physical system or may substantially reduce the attainable set of models. Zero elements and partial dependences constitute "minor restrictions", which may make some equation structures unattainable, requiring redefinition of the target structure.

A simple example is given below illustrating the basic concept of transformation.

**Example 2.** To illustrate the transformation procedure, consider the following static system without plant disturbance faults:

$$
b_{11} \Delta u_1(t) + b_{12} \Delta u_2(t) = y_1(t)
$$

$$
b_{21} \Delta u_1(t) + b_{22} \Delta u_2(t) = y_2(t).
$$

The parity equations (without noise) are:

$$
b_{11} \Delta y_1(t) + b_{12} \Delta y_2(t) - \Delta y_1(t) = e_1(t)
$$

$$
b_{21} \Delta y_1(t) + b_{22} \Delta y_2(t) - \Delta y_2(t) = e_2(t).
$$

The incidence matrix

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

indicates that the system is not isolable even with zero thresholds. By eliminating, for example, $\Delta u_1(t)$ from the two equations we obtain

$$
\begin{bmatrix}
b_{12} & b_{11} \\
- b_{22} & b_{21}
\end{bmatrix} \Delta y_2(t) - \Delta y_1(t) + \begin{bmatrix}
b_{11} \\
- b_{12}
\end{bmatrix} \Delta y_1(t) = e_1(t).
$$

The incidence matrix for the three equations is

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

indicating that the system is now isolable with zero threshold but not with high threshold. Eliminating $\Delta u_2(t)$ from the two primary equations yields

$$
\begin{bmatrix}
b_{11} & b_{12} \\
- b_{22} & b_{21}
\end{bmatrix} \Delta u_1(t) - \Delta y_1(t) + \begin{bmatrix}
b_{12} \\
- b_{11}
\end{bmatrix} \Delta y_2(t) = e_1(t).
$$

The incidence matrix for the full transformed system becomes

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

column-canonical (of order 1), indicating that statistical (high-threshold) isolability has been attained.

**Transformation procedure.** According to equation (8), the primary residuals $e(t)$ depend on the additive faults $\nu(t)$ and $\Delta x(t)$ as

$$
e(t) = -G(z)\nu(t) + \hat{F}(z)\Delta x(t)
$$

$$
= [G(z) \quad \hat{F}(z)] \begin{bmatrix}
-\nu(t) \\
\Delta x(t)
\end{bmatrix}.
$$

(13)

The new residuals $e^*(t)$ are sought as a linear transformation of the primary ones:

$$
e^*(t) = C(z)e(t).
$$

(14)

The transformed residuals will depend on the additive faults via the transformed matrices $G^*(z)$ and $F^*(z)$:

$$
e^*(t) = [G^*(z) \quad F^*(z)] \begin{bmatrix}
-\nu(t) \\
\Delta x(t)
\end{bmatrix}.
$$

(15)

That is,

$$
[G^*(z) \quad F^*(z)] = C(z)G(z) \hat{F}(z).
$$

(16)
The target system \([G^*(z) \ F^*(z)]\) is specified in terms of its incidence matrix \([\Gamma^* \ \Phi^*]\). Elements of the transforming matrix \(C(z)\) are defined by the position of zeroes in the target set and by the numerical values of elements in the primary set. The resulting transformation yields numerical values for the target matrices, with the desired structure.

Each row of equation (16) can be handled separately. The \(i\)th row of \(C(z)\), \(c_i(z)\) is related to the \(i\)th row of \([G^*(z) \ F^*(z)]\), each zero specifying a condition (a homogeneous linear equation) for the elements \(c_i(z)\) of \(c_i(z)\).

Since \(c_i(z)\) has \(m\) elements, the maximum number of zeroes in each row of the target set is \(m-1\), with one of the elements of \(c_i(z)\) being free.

A zero in the \(i\)th position of the \(\Gamma^*\) matrix yields the homogeneous linear equation

\[
g_{ji}(z) = \sum_{i=1}^{m} c_i(z)g_i(z) = 0 \tag{17}
\]

while a zero in the \(j\)th position of \(\Phi^*\) yields

\[
f_{ji}(z) = \sum_{i=1}^{m} c_i(z)f_i(z) = 0. \tag{18}
\]

In the latter case, if \(j > k\) [that is, \(f_{ij}(z)\) is in the \(A^*\) part of \(\Phi^*\)], and from the primary model \([G(z) \ F(z)]\), row-by-row, in the following steps:

• from the zeroes in the \(A^*\) part of \(\Phi^*\) (last \(m\) columns), obtain the \(c_i(z)\) = 0 elements according to equation (19);

• place and assign values to the \(m-r_i>0\) free elements, where \(r_i\) is the number of zeroes in the concerned row of \([\Gamma^* \ \Phi^*]\);

• find the remaining elements by solving the simultaneous equations obtained according to (17) and (18);

• multiply the entire row with the common denominator of its elements.

Example 3. To illustrate the transformation procedure, let us consider the following system: \(m = 4, \ k = 2, \ q = 0\)

\[
F = \begin{bmatrix}
 b_{11} & b_{12} & 1 & 0 & 0 & 0 \\
 b_{21} & b_{22} & 0 & 1 & 0 & 0 \\
 b_{31} & b_{32} & 0 & 0 & 1 & 0 \\
 b_{41} & b_{42} & 0 & 0 & 0 & 1 \\
\end{bmatrix}
\]

with the target set

\[
\Phi^* = \begin{bmatrix}
 1 & 1 & 0 & 0 & 0 \\
 0 & 1 & 0 & 1 & 0 \\
 0 & 0 & 1 & 0 & 1 \\
\end{bmatrix}
\]

The transforming matrix \(C\) is \(4 \times 4\). According to equation (19), \(C_{ij} = 0\) whenever the \(A^*\) part (the last \(m\) columns) of \(\Phi^*\) contains a zero at the respective position, yielding

\[
C = \begin{bmatrix}
 c_{11} & 0 & 0 & 0 \\
 0 & c_{22} & c_{23} & 0 \\
 0 & c_{32} & c_{33} & 0 \\
 c_{41} & 0 & c_{43} & c_{44} \\
\end{bmatrix}
\]

The remaining elements can be obtained row by row. First row:

\(c_{11}\) is free, e.g. \(c_{11} = 1\).

Second row:

\(c_{22}b_{22} + c_{24}b_{42} = 0\)

one of \(c_{22}, \ c_{24}\) is free, e.g. \(c_{22} = 1\),

\(c_{24} = -\frac{b_{42}}{b_{22}}\)

after multiplication with denominator:

\(c_{22} = b_{22}, \ c_{24} = -\frac{b_{42}}{b_{22}}\).

Third row:

\(c_{32}b_{21} + c_{33}b_{31} = 0\)

one of \(c_{32}, \ c_{33}\) is free, e.g. \(c_{32} = 1\),

\(c_{33} = -\frac{b_{31}}{b_{21}}\)

after multiplication with denominator:

\(c_{32} = b_{31}, \ c_{33} = -\frac{b_{31}}{b_{21}}\).

Fourth row:

\(c_{41}b_{11} + c_{42}b_{21} + c_{43}b_{31} = 0\)

one of \(c_{41}, \ c_{42}, \ c_{43}\) is free, e.g. \(c_{41} = 1\);

with this

\(c_{43}b_{31} + c_{44}b_{41} = -b_{11}\)

\(c_{43}b_{31} + c_{44}b_{41} = b_{12}\)

yielding

\(c_{43} = \frac{b_{12}b_{41} - b_{11}b_{42}}{b_{31}b_{42} - b_{32}b_{41}}\)

\(c_{44} = \frac{b_{13}b_{32} - b_{12}b_{33}}{b_{31}b_{42} - b_{32}b_{41}}\)

after multiplication with common denominator:

\(c_{41} = b_{31}b_{42} - b_{32}b_{41}, \ c_{43} = b_{12}b_{41} - b_{11}b_{42}, \ c_{44} = b_{13}b_{32} - b_{12}b_{33}\).

Note that the above model generation procedure may be interpreted as a special case of the more general technique proposed by Chow and Willsky (1984). Based on the state-space model of the system, Chow and Willsky generated arbitrary parity equations by manipulating on a generalized observability matrix. This latter procedure will also yield the parity equations required for isolability if the appropriate structural constraints (zero coefficients) are imposed on it.

Restrictions on the transformations. Substantial deficiencies of the input part \([G(z) \ B(z)]\) of the system matrix, such as all zero columns, full column dependence or full row dependence, represent "major restrictions", not only in terms of model transformation, but also in overall isolability and detectability of certain failures. Although these seem rather extreme situations, they appear frequently in real-life systems. In most cases, they cannot be cured by algebraic manipulations but may require alterations of the physical system.

Major restrictions of the system matrix \([G(z) \ B(z)]\) have the following consequences:

(a) If a column contains all zero elements, the respective failure is not detectable and cannot be made detectable by any transformation.

(b) If there is full linear dependence among \(k\) columns, then removing any \(k-1\) of the concerned failures eliminates the \(k\)th failure as well. In particular, if \(k = 2\), the two failures may only be removed together; that is, they cannot be made isolable.

(c) If there are \(\mu\) linear relationships among the \(m\) full rows, not more than \(m-1-\mu\) failures may be removed without eliminating all the other failures as well.

An all-zero column signifies that all primary residuals are orthogonal to the concerned failure. This property transfers to any transformed model. That is, the failure is and remains undetectable. This can only be cured by adding, if possible,
new equations to the model, with new output variables, that are not orthogonal to the concerned input; this implies additional sensors in the plant.

If \( k \) columns are completely linearly dependent, then the projection of all residuals on the subspace of the concerned failures is constrained to a \( k - 1 \) dimensional sub-subspace. This means that any residual may be made orthogonal to \( k - 1 \) of the concerned failures only by making the projection zero. If \( k = 2 \), such failures are not and cannot be made isolable. Full column dependence with \( k = 2 \) happens, for example, if the sum of two or more input variables enters a system. This problem can only be cured by adding one or more redundant measurements (sensors) to the physical system.

The projections of the \( m \) primary residuals on the subspace of all inputs, in general, span an \( m \)-dimensional sub-subspace. If there are \( m \) linear relationships among the projections (\( m \) full row dependences), then the order of the sub-subspace is reduced to \( m - \mu \). Any linear combination of these residuals is confined to the same sub-subspace; orthogonality to \( m - \mu \) or more failures may be achieved only if the projection becomes zero. Full row dependence occurs, for example, when output variables are the outputs of cascaded units with no intermediate inputs. Full row dependence does not necessitate any alteration of the physical system but may significantly constrain the attainable set of transformed models.

**Example 4.** This example illustrates the consequences of full column and row dependence and some of the possible remedies. Consider the system shown in Fig. 1, without the redundant measurement \( \hat{y}_3(t) \). The input-output equations are

\[
\begin{align*}
&u_1(t) + b_1u_2(t) = (1 - d_1z^{-1})y_1(t) \\
&u_2(t) + b_2u_2(t) = (1 - d_2z^{-1})(1 - d_2z^{-1})y_2(t).
\end{align*}
\]

The corresponding parity equations are

\[
\begin{align*}
&b_1\Delta u_1(t) + b_2\Delta u_2(t) - a_2(z)\Delta y_1(t) = e_1(t) \\
&b_2\Delta u_2(t) + b_2\Delta u_2(t) - a_2(z)\Delta y_2(t) = e_2(t),
\end{align*}
\]

with \( a_2(z) = 1 - d_1z^{-1} \) and \( a_2(z) = 1 - (d_1 + d_2)z^{-1} + d_1d_2z^{-2} \). The incidence matrix is

\[
\Phi = \begin{bmatrix}
0 & 1 & 1 & 1 \\
1 & 1 & 0 & 1 \\
1 & 1 & 1 & 0
\end{bmatrix}
\]

showing the system is not isolable. Also, the \( \mathbf{B} \) matrix has full column and row dependence.

The full column dependence can only be remedied by adding a redundant measurement \( \hat{y}_3(t) \) (a second sensor measuring \( u_1 \); see Fig. 1). This introduces a third parity equation

\[
\Delta u_1(t) - \Delta y_3(t) = e_3(t)
\]

and the extended incidence matrix

\[
\Phi^+ = \begin{bmatrix}
1 & 1 & 1 & 0 \\
1 & 1 & 0 & 1 \\
1 & 0 & 0 & 1
\end{bmatrix}
\]

The system is now zero-threshold isolable. To achieve high-threshold isolability, model transformation needs to be applied. The full row dependence between the first two equations and a zero on the input side in the third restrict the attainable column-canonical models. It appears that the only attainable column-canonical structure is

\[
\begin{bmatrix}
0 & 1 & 1 & 1 \\
1 & 0 & 1 & 1 \\
1 & 1 & 0 & 1 \\
1 & 1 & 1 & 0
\end{bmatrix}
\]

A suitable transforming matrix is found as

\[
\begin{bmatrix}
1 & 1 & -2b_1 \\
1 & -1 & 1 \\
1 & 0 & 1 \\
1 & 1 & 0
\end{bmatrix}
\]

Zero elements and partial dependence in the \([G(z) \hat{B}(z)]\) matrix place “minor restrictions” on the transformation procedure. Partial dependence is a situation when any part (two or more elements) of a row or column of the \([G(z) \hat{B}(z)]\) matrix is linearly dependent on the respective part of other row(s) or column(s). As a consequence of minor restrictions, some of the target equations may not be attainable; the transformation then either aborts or leads to structures different from the target set.

- (a) simultaneous equations of types (17) and (18) may be unsolvable;
- (b) simultaneous equations of types (17) and (18) may yield zero values for some of the elements of matrix \( \mathbf{C}(z) \) in addition to the ones obtained from the single equations (19). Such zero elements produce excess zeroes in the \( \mathbf{A}^*(z) \) part (last \( m \) columns) of the transformed system ("explicit excess zeroes"); thus, it is different from the target set;
- (c) some rows of matrix \( \mathbf{C}(z) \), combined with columns of \( \mathbf{G}(z) \) or \( \hat{\mathbf{B}}(z) \), produce excess zeroes in the \( \mathbf{G}^*(z) \) or \( \hat{\mathbf{B}}^*(z) \) part of the transformed system. This cannot be seen from \( \mathbf{C}(z) \) and can only be detected by performing the transformation ("implicit excess zeroes").

If the selected target set proves unattainable, another one is to be chosen heuristically, or based on sensitivity/robustness considerations. This is done without difficulty since the number of possible sets is very large, as will be shown later in this paper.

**Design issues**

Designing a failure detection and isolation system, within the framework of high-threshold isolability for additive failures, implies the following steps:

- Select the overall model structure, that is, the number of rows in the model and the number of zeroes per column.
- Select the particular column-canonical model structure, that is, assign the position of zeroes.
• Obtain numerical values for the parameters.

Within the selected overall model structure, there are usually a very large (combinatorial) number of particular model structures that satisfy the isolability conditions. This situation, which will be referred to as model redundancy, is a mixed blessing. On the one hand, it makes model selection very difficult and, in most cases, intractable for the usual mathematical tools. On the other hand, it allows the introduction of other, non-structural considerations into the design procedure, most notably the aspect of robustness with respect to modeling errors. The numerical values of the model parameters are either unique for the given row structure or represent additional degrees of freedom in the pursuit of robustness.

**Model redundancy.** In choosing the overall model structure, one should strive for the minimum number of rows (to keep the algorithm simple) and for the maximum number of zeroes per column (to enhance isolability). The choice must guarantee that there is a sufficient number of different column structures; that is, the following inequality must be satisfied:

$$\left( \frac{p}{r} \right) \geq k + q + m$$

(20)

where

- $p$ is the number of rows
- $r$ is the number of zeroes per column

and $k$, $q$ and $m$ are the number of the measured inputs, disturbances and outputs, respectively. Some additional guidance is provided by another inequality, reflecting the fact that no more than $m - 1$ zeroes per row may be specified for the model transformation. Thus, the overall balance of zeroes requires that

$$r(k + q + m) \leq p(m - 1).$$

(21)

Satisfaction of the overall balance, however, does not guarantee that all equations within the target model are attainable. Note that the number of rows $p$ that is needed to satisfy both inequalities is never greater than $k + q + m$ and, in many cases, is even smaller than $m$.

For the set of possible column-canonical structures is of combinatorial size. Many of these structures may contain unattainable equations and thus are not really useful. Others may be distinct column-wise but are composed of the same set of parity equations, thus are not really different models. Still the remaining set of attainable and different models is extremely large; depending on the number of variables in the system, they may run in the thousands and beyond.

The number of model choices may be reduced if we restrict ourselves to parity equations row-canonically of order $m - 1$. Equations of the primary model are normally of this kind; secondary equations may be obtained by transformation. These secondary equations have unique parameters. The full set of equations row-canonically of order $m - 1$

$$P = \frac{k + q + m}{m - 1}$$

(22)

Part of this set may not be attainable, owing to restrictions (zeroes and dependences) in the primary model. Although the number of row-canononical equations is significant, an enumeration and analysis of the full attainable set may be feasible, or a further reduction of the choices results if the structures are restricted to (redundant) square models; that is, $p = k + q + m$.

**Design for isolability and robustness.** Although the methodology described in this paper does not address the modeling error problem explicitly, model redundancy offers a way to incorporate robustness considerations into the design. A numerical measure of robustness may be defined and computed for each parity equation (or its free parameters) optimized with respect to such measure). The best model is sought then, in principle, as the one providing optimum robustness under the structural constraint posed by isolability.

A possible measure of robustness is the limit value of a selected modeling error that triggers the threshold test on a given residual, under no failure and zero momentary noise (Gertler, 1986d). The modeling error is related to a specific parameter or group of parameters, usually in an underlying model, such as the state-space model or a problem-specific plant representation (e.g., a time constant, a heat-transfer coefficient, etc.). A difficulty of such robustness measures is that they also depend on the plant operating point and the noise statistics.

Finding the model that satisfies the isolability conditions and is optimal (or at least acceptable) from the point of view of robustness is generally a tedious trial-and-error procedure. After determining the overall model parameters $p$ and $r$, an arbitrary column-canonical structure can be chosen, then model transformation attempted. If the structure proves unattainable, a new structure has to be chosen heuristically. Once an attainable structure has been found, the selected robustness measure is computed for each parity equation. In the case of unsatisfactory robustness, the model may be improved by replacing its rows with other equations, one at a time, starting with the poorest performer. However, replacement of each row should be accompanied by at least one other replacement, in order to maintain the column-canonical structure of the model. If the model is restricted to row-canonical equations and square structure, the above search can be automated and leads to a reasonably robust set of equations within the isolable structure. This search algorithm will be reported in detail elsewhere.

The design procedure may be aimed at obtaining a single model, "best" for a selected single robustness measure, or a bank of models, each designed with a different robustness requirement in mind. These latter models can then be used in parallel and are certainly helpful, for example, when some of the failure signatures are indecisive, owing to the marginal size of the failure. While model generation is normally an activity performed off-line, the idea may be developed further to include a varying pool of models generated on-line, in accordance with the momentary failure/noise/modeling error situation.

**Conclusion**

A new framework has been described for developing parity equations that prevent incorrect isolation decisions under marginal size failures, provided each residual is tested independently and the thresholds are set high. The approach utilizes residuals that are orthogonal (to certain failures and arranges these into special isolable structures. A transformation algorithm provides a multitude of models that satisfy the isolability requirements. A search procedure, utilizing this model redundancy, integrates model error robustness considerations into the design. The method can be applied to single additive-type failures on the measured input and output variables and to simple additive plant disturbances, provided the input–output model and the disturbance model are known.

Further research is required to extend the method to multiplicative-type failures and multiple failures. The efficiency of the model search procedure needs to be significantly improved in the general case, when the model is not restricted to a square structure of canonical rows. Also some technique is to be developed to handle incomplete and/or contradictory isolation decisions, especially when several models are used in parallel.

**References**


