Robust Isolable Models for Failure Diagnosis

The paper presents a design procedure to generate parity equation models for analytical failure detection and isolation. The models are in "statistically-isolable" structure that prevents incorrect isolation decisions under marginal failure size if the residuals are tested in parallel. Of the combinatorial multitude of structurally-suitable models, a choice is made on the basis of sensitivities with respect to different failures and robustness relative to uncertainties in selected parameters of an underlying model. A complete set of parity equations fitting the isolable structure is first generated from a state-space model. The equations unsuitable for sensitivity are eliminated, and a suboptimally-robust diagnostic model is obtained in a constrained discrete search over the remaining set.

Janos Gertler
Qiang Luo
School of Information Technology and Engineering
George Mason University
Fairfax, VA 22030

Introduction

The detection and diagnosis of failures in the plant and in its instrumentation have always been one of the main functions of computerized plant monitoring systems. While most of the implemented systems still rely on checking plant measurements against fixed limits, several more advanced methods have also been developed. These include mulitple sensors, special sensors (such as limit sensors, noise, and vibration) and the frequency analysis of plant signals. If a reasonably accurate model of the plant is available or if it can be obtained by identification, this model may be utilized in analytical failure detection techniques.

A significant class of analytical failure detection and isolation (diagnostic) methods is based on the input-output model of the system. In the traditional chemical engineering practice, balance equations have been used for this purpose; these are usually static and restricted to extensive variables (mass and energy). A more general approach utilizes the complete dynamic system model. Such "generalized" balance equations are also called parity equations. These and other analytical techniques have been recently surveyed by Gertler (1988).

The value a parity equation returns, called residual, is ideally zero. With real data, the residuals are usually nonzero, due to noise, modeling errors and occasional failures in the plant, actuators or sensors. To account for the effect of random noise, failure decisions are usually based on statistical tests applied to the residuals. Such testing may be performed sequentially using some aggregate statistic or in parallel, separately for each residual. In the latter case, special model structures may enhance failure isolation.

Vaclavek (1974) proposed parallel statistical tests on the short-term averages of balance equation residuals. Almasy and Sztano (1975) detected bias faults on the basis of the size of elements in a trnasformed residual vector. Iordache et al. (1985) applied parallel tests to measurement residuals following data reconciliation. A geometric theory of static parity equations was presented by Potter and Suman (1977). Ben Haim (1980) used redundant balance equations for residual generation. Chow and Willsky (1984) proposed a systematic procedure to generate parity equations from the state-space representation of a dynamic system. Sequential algorithms were proposed by Mah et al. (1976), Romagnoli and Stephanopoulos (1981), and Crowe (1988). Kramer (1987) integrated balance models into diagnostic expert systems, replacing boolean test results with smoothed belief functions.

A fundamental issue of diagnostic algorithms is isolability, that is, the ability of the algorithm to distinguish different failures. If the residuals are investigated separately, in parallel, then this isolability property requires residuals that are orthogonal to certain failures. This places a constraint on the structure of the set of parity equations. Ben Haim (1980) formulated a structural requirement along these lines, with the implicit assumption that the test thresholds are close to zero ("deterministic isolability"). He also suggested (1983) a model transformation to attain such structures for dynamic systems. A more general isolability condition, that prevents incorrect isolation decisions under marginal size failures ("statistical isolability"), was introduced by Gertler and Singer (1985, 1990). A transformation technique was also presented that, when applied to the original input-output model of the plant, produces the desired parity equations.

The sensitivity of parity equations to certain failures is an important consideration in the design of the diagnostic system.

While sensitivities may be controlled by the appropriate filtering of the residuals (Gertler and Singer, 1985), their ratio within a given parity equation remains constant. To characterize the quality of parity equations from this point of view, the concept of "sensitivity condition" will be introduced in this paper.

While static balance equations are free of modeling errors, this is usually not the case if dynamic system models are concerned. The inevitable errors in most dynamic models may seriously interfere with the detection and isolation of failures. Therefore, robustness relative to such errors is a critical issue in the design of diagnostic algorithms. Chow and Willsky (1984) formulated a conceptual minimax framework for robust design. Subsequently, Lou et al. (1986) developed a design methodology that provides the most robust parity equations given a finite set of uncertain plant models. As an alternative and perhaps more practical approach, the concept and measure of "partial robustness" was introduced (Gertler, 1986).

The subject of the present paper is a complete design procedure that generates the "most robust" parity equation model within the structure needed for statistical isolability while simultaneously satisfying certain minimum sensitivity condition requirements. The individual equations are obtained from a state-space description of the plant by the Chow-Willsky technique, modified to provide for orthogonality with respect to selected failures. The sensitivity condition is computed for each equation, together with the partial robustness measures relative to uncertainty in selected underlying model parameters. The equations not meeting the required sensitivity ratio are eliminated from the set. The most robust model is then found in a constrained discrete search over the reduced set of equations. For the search procedure, both a global and local variant have been developed.

The complete procedure has been computer-implemented (Luo, 1989). Its operation will be demonstrated on a design example using a linearized dynamic model of a nine-plate distillation column (Takamatsu et al., 1979).

Problem Description

Consider a linear (linearized) plant with inputs $\boldsymbol{u}(t) = [u_1(t), \ldots, u_k(t)]^T$ and outputs $\boldsymbol{y}(t) = [y_1(t), \ldots, y_m(t)]^T$. The plant is described by a discretized model. It is assumed that the plant parameters are constant. A state-space model can be written as

$$\begin{aligned} \boldsymbol{x}(t) &= \boldsymbol{A}\boldsymbol{x}(t-1) + \boldsymbol{B}\boldsymbol{u}(t-1) + \boldsymbol{L}\boldsymbol{w}(t-1) \\ \boldsymbol{y}(t) &= \boldsymbol{C}\boldsymbol{x}(t) + \boldsymbol{D}\boldsymbol{u}(t) \end{aligned} \tag{1}$$

Here $\boldsymbol{x}(t) = [x_1(t), \ldots, x_n(t)]^T$ is the state vector, $\boldsymbol{w}(t)$ is the plant noise, and $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D}, \boldsymbol{L}$ are system matrices of appropriate size.

Introduce the backward shift operator z^{-1}. With this, the state equation can be rewritten as

$$\begin{aligned} \boldsymbol{x}(t) &= \boldsymbol{A}z^{-1}\boldsymbol{x}(t) + \boldsymbol{B}z^{-1}\boldsymbol{u}(t) + \boldsymbol{L}z^{-1}\boldsymbol{w}(t) \\ \boldsymbol{y}(t) &= \boldsymbol{C}\boldsymbol{x}(t) + \boldsymbol{D}\boldsymbol{u}(t) \end{aligned} \tag{2}$$

Solving for the output yields

$$\begin{aligned} \boldsymbol{y}(t) &= \boldsymbol{C}(\boldsymbol{I} - z^{-1}\boldsymbol{A})^{-1}[z^{-1}\boldsymbol{B}\boldsymbol{u}(t) + z^{-1}\boldsymbol{L}\boldsymbol{w}(t)] + \boldsymbol{D}\boldsymbol{u}(t) \\ &= \frac{\boldsymbol{G}(z^{-1})}{h(z^{-1})}\boldsymbol{u}(t) + \frac{\boldsymbol{E}(z^{-1})}{h(z^{-1})}\boldsymbol{w}(t) \end{aligned} \tag{3}$$

where $\boldsymbol{G}(z^{-1})$ and $\boldsymbol{E}(z^{-1})$ are matrix polynomials and $h(z^{-1})$ is a scalar polynomial in the shift operator:

$$\begin{aligned} \boldsymbol{G}(z^{-1}) &= \boldsymbol{G}_0 + \boldsymbol{G}_1 z^{-1} + \cdots + \boldsymbol{G}_n z^{-n} \\ \boldsymbol{E}(z^{-1}) &= \boldsymbol{E}_1 z^{-1} + \cdots + \boldsymbol{E}_n z^{-n} \\ h(z^{-1}) &= 1 + h_1 z^{-1} + \cdots + h_n z^{-n} \end{aligned} \tag{4}$$

Equation 3 can also be written as

$$\boldsymbol{H}(z^{-1})\boldsymbol{y}(t) = \boldsymbol{G}(z^{-1})\boldsymbol{u}(t) + \boldsymbol{E}(z^{-1})\boldsymbol{w}(t) \tag{5}$$

Here $\boldsymbol{H}(z^{-1})$ is a diagonal matrix.

The inputs $\boldsymbol{u}(t)$ may be subdivided into three groups:

1. Measured inputs $\boldsymbol{u}_M(t)$. These are not controlled (at least locally) but are measured using some sensor (Figure 1). Their

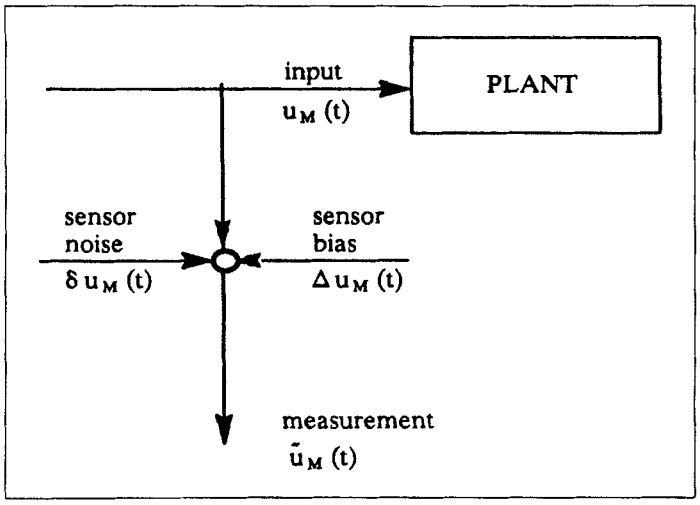

Figure 1. Measured input.

measured value $\tilde{u}_M(t)$ is subject to measurement noise $\delta u_M(t)$ and, potentially, a bias type sensor fault $\Delta u_M(t)$:

$$\tilde{u}_M(t) = u_M(t) + \delta u_M(t) + \Delta u_M(t) \tag{6}$$

2. Controlled inputs $u_C(t)$. These are controlled, usually by some actuator, but not measured. Now $\tilde{u}_C(t)$ is the intended (computed) input value and $u_C(t)$ is its realization by the actuator. The actuator noise $\delta u_C(t)$ and any potential actuator fault $\Delta u_C(t)$ are transformed on the actuator input to facilitate uniform treatment with the measured inputs (Figure 2):

$$\tilde{u}_C(t) = u_C(t) + \delta u_C(t) + \Delta u_C(t) \tag{7}$$

3. Disturbance inputs $u_D(t)$. These are neither measured nor controlled. Normally they are zero; any nonzero value may be considered as a plant fault.

The outputs $y(t)$ are all measured; their measured value $\tilde{y}(t)$ is subject to measurement noise $\delta y(t)$ and, potentially, bias type sensor fault $\Delta y(t)$

$$\tilde{y}(t) = y(t) + \delta y(t) + \Delta y(t) \tag{8}$$

Any input-output relationship qualifies as a parity equation (generalized balance equation), provided that it contains only accessible (measured or controlled) variables and is satisfied under ideal (noise- and failure-free) situations. Equation 5, with the plant disturbances $u_D(t)$ and plant noise $w(t)$ omitted, provides a set of m linearly independent such equations:

$$H(z^{-1})y(t) = G_M(z^{-1})u_M(t) + G_C(z^{-1})u_C(t) \tag{9}$$

Where $G_M(z^{-1})$ and $G_C(z^{-1})$ are the appropriate submatrices of $G(z^{-1})$.

Practically, this equation can only be computed with the available quantities $\tilde{u}_M(t)$, $\tilde{u}_C(t)$, $\tilde{y}(t)$ and usually will not be exactly satisfied, due to the noises and potential faults. The equation error or residual $r(t)$ is obtained as

$$r(t) = [G_M(z^{-1}), G_C(z^{-1}), -H(z^{-1})] \begin{bmatrix} \tilde{u}_M(t) \\ \tilde{u}_C(t) \\ \tilde{y}(t) \end{bmatrix} = F(z^{-1})\tilde{q}(t) \tag{10}$$

Here $q(t)$ denotes the combined measurement vector (including the controlled inputs) and $F(z^{-1})$ the combined coefficient matrix. Equation 10 is the "external" or computational form of the parity equations. To see how the residuals depend on the noises and faults, substitute Eq. 5 through 8 into the external equation. This leads to the following "internal" form:

$$r(t) = F(z^{-1})\delta q(t) - E(z^{-1})w(t) + F(z^{-1})\Delta q(t) - G_D(z^{-1})u_D(t) \tag{11}$$

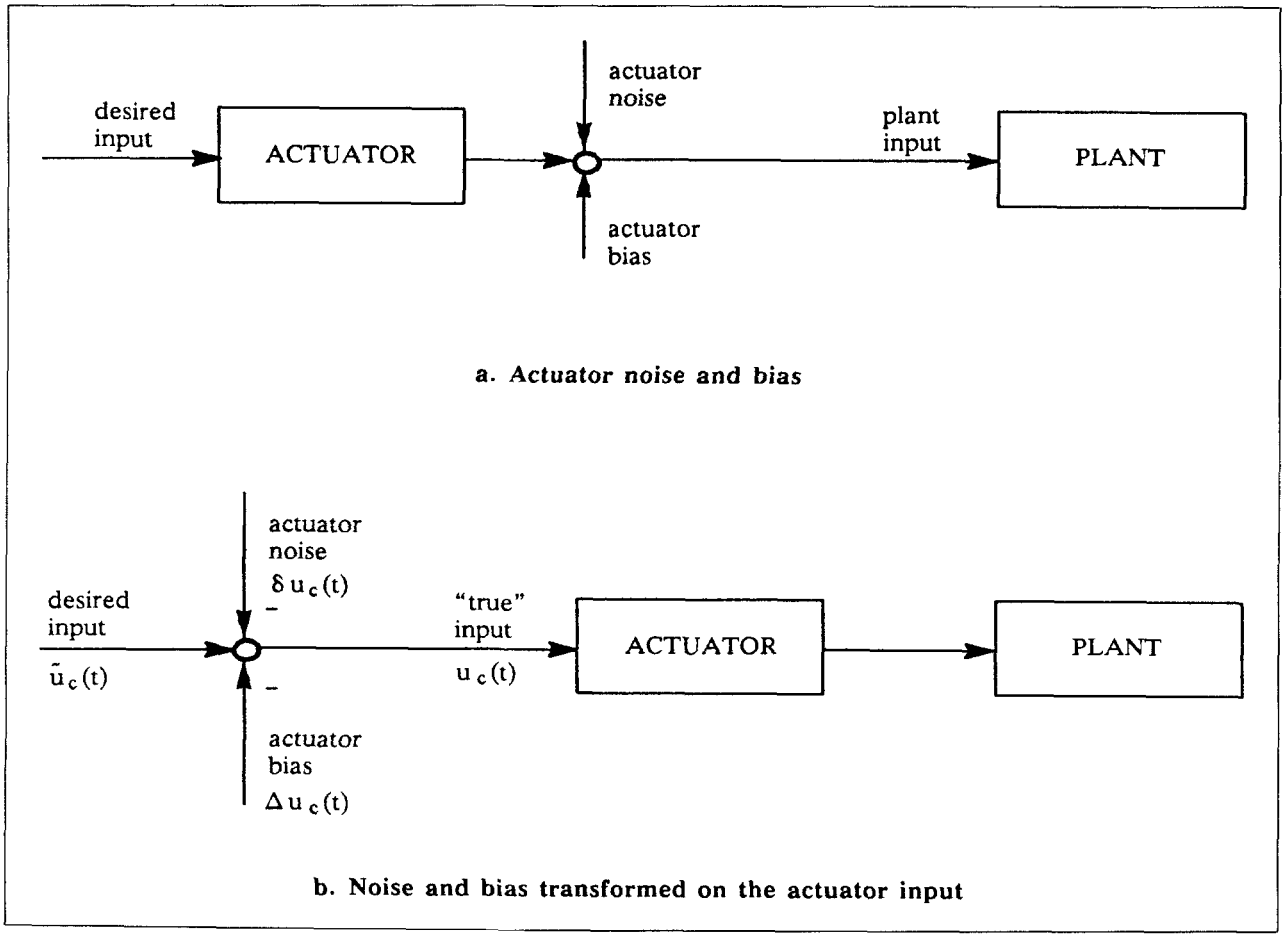

Figure 2. Controlled input.

Here $\delta q(t)$ and $\Delta q(t)$ are the noise and faults, respectively, associated with the combined vector $q(t)$.

Note that any discrepancy between the plant model and the true plant parameters would also contribute to the residuals. Such discrepancies may be due to multiplicative (parametric) plant faults or to modeling errors. The former will not be treated in this paper. Modeling errors will be discussed in a later section.

Statistical Testing

As pointed out earlier, the residuals are subjected to statistical testing, in order to separate failures (that, once appeared, are assumed to stay) from the random effects of noise. While other testing approaches (like forming an aggregate scalar statistic over the vector of residuals) are also possible, in the methodology discussed here the residuals are tested individually, in parallel. Besides the obvious simplicity of this approach, our primary motivation has been the natural link from parallel testing to parallel fault isolation.

To establish the test thresholds, consider Eq. 11, with the assumption that there is no failure in the system

$$r(t) = F(z^{-1})\delta q(t) - E(z^{-1})w(t) \tag{12}$$

Now $r(t)$ is a random vector, determined by the noises $\delta q(t)$ and $w(t)$ via a discrete moving average relationship. If the noises are normally distributed with zero mean, so are the residuals. If the noise variances are known (or guessed) the variances of the residuals can be computed. Assuming that the noises are white and $\delta q(t)$ is independent of $w(t)$, the zero-shift covariance matrix S_r of the residual vector is obtained as

$$S_r = \sum_{l=0}^{n} [F_l S_q F_l^T + E_l S_w E_l^T] \tag{13}$$

where S_q and S_w are the (zero shift) covariance matrices of $\delta q(t)$ and $w(t)$, respectively. Furthermore, if S_q and S_w are diagonal, then the individual residual variances may be computed as

$$\text{var}\{r_i(t)\} = \sum_{j=l}^{k+m}\left[\sum_{l=0}^{n} f_{lij}^2\right]\text{var}\{\delta q_j(t)\} + \sum_j\left[\sum_{l=0}^{n} e_{lij}^2\right]\text{var}\{w_j(t)\} \tag{14}$$

where f_{lij} and e_{lij} are the ijth element of F_l and E_l, respectively.

The test thresholds are defined as a multiple of the residual's standard deviation. To avoid frequent false alarms, usually high thresholds are selected.

The test thresholds may be reduced, without increasing the frequency of false alarms, by filtering the residuals (Gertler and Singer, 1985). A separate and potentially different filter may be applied to each residual. The filter can be designed in such a way that the variance of the filtered residual is much smaller than the variance of the original one while their means (the effect of an occasional failure) are the same.

Consider, for example, a second-order discrete filter:

$$\Gamma(z^{-1}) = \frac{1 + \gamma_1 + \gamma_2}{1 + \gamma_1 z^{-1} + \gamma_2 z^{-2}} \tag{15}$$

The shown choice of the numerator guarantees that the filter has unit gain with respect to the mean. By selecting a particular filter structure (e.g., Butterworth filter) all filter parameters become determined by a single parameter. A formula to compute the variance of the filtered residual, under the assumption of zero mean, whiteness and mutual independence of the noises, is given in the Appendix.

Note that filtering, while lowering the thresholds, also introduces a delay in detection. In many cases, however, this delay is only apparent since detection would, in fact, never happen without lowering the threshold.

Sensitivity

While the test thresholds for each parity equation are determined by the noise with no regard to the failure sizes, the question arises what size of failures will then drive the residuals to their respective thresholds. Since in the presence of a failure any residual is the combined effect of this failure and noise, it is more accurate to ask what size of failure will drive the mean of the residual's distribution to the threshold or, equivalently, what is the critical size of a failure that results in a residual equal to its threshold if the instantaneous value of the combined noise is zero.

A sensitivity measure along these lines, called "triggering limit," was introduced in earlier work (Gertler, 1986). Obviously, the triggering limit depends on the threshold and on the "gain" of the parity equation relative to the failure. In a dynamic environment, further clarifications are necessary: one may assume that the failure acts as a step function and then the gain may be defined by its steady-state value or maximum transient. This leads, for example, to the following definition for the triggering limit ρ_{ij}:

$$\rho_{ij} = \frac{\eta_i}{|f_{ij}(z^{-1})|} \tag{16}$$

Here η_i is the threshold for the ith residual and $|f_{ij}(z^{-1})|$ denotes the steady-state gain belonging to the jth failure in the ith equation. Note that the triggering limit, in general, is different for each failure within a given parity equation and also in each equation for a given failure.

An alternative and probably more practical sensitivity measure can be obtained by defining, as part of the problem specification, the nominal value Δq_j^* for each failure variable that should, under zero instantaneous noise, bring the given residual to its threshold. Assuming again a step function failure, a steady-state sensitivity number λ_{ij} may be computed as

$$\lambda_{ij} = \frac{\eta_i}{\Delta q_j^* |f_{ij}(z^{-1})|} \tag{17}$$

Ideally, each sensitivity number should be 1. If the threshold is adjusted by filtering, the sensitivity numbers change accordingly. The necessary adjustment for any given equation may be determined by its maximum or average sensitivity number.

Filtering does not influence the ratio of sensitivity numbers within a parity equation. The maximum ratio will be called the "sensitivity condition" κ_i of the equation

$$\kappa_i = \frac{\max_j \{\lambda_{ij}\}}{\min_j \{\lambda_{ij}\}} = \frac{\max_j \{\Delta q_j^* |f_{ij}(z^{-1})|\}}{\min_j \{\Delta q_j^* |f_{ij}(z^{-1})|\}} \tag{18}$$

The sensitivity condition is the most important sensitivity measure of a parity equation. Ideally it should be 1; if it is much larger, then there is an obvious imbalance among the different failures in the equation. If the threshold is adjusted, by filtering, to satisfy the failure with the smallest sensitivity number, failures with larger sensitivity numbers have to exceed their nominal value significantly to trigger the test. On the other hand, if the threshold is adjusted to the largest sensitivity number, failures with smaller sensitivity numbers will trigger the test even if their size is a fraction of the nominal value, causing false alarms and seriously hindering failure isolation. Since the sensitivity condition of any given parity equation cannot be influenced, equations with excessive condition measure should be identified and eliminated from the design.

Modeling Error Robustness

Dynamic plant models are inevitably subject to modeling errors. This may be due to linearization, lower-order approximation of higher-order systems or simply parameter errors. Modeling errors interfere with model-based failure detection and isolation in that they contribute some additional terms to the residuals.

Assume the model $\hat{\boldsymbol{F}}(z^{-1})$ describes the true plant $\boldsymbol{F}(z^{-1})$ with a modeling error $\Delta\boldsymbol{F}(z^{-1})$:

$$\hat{\boldsymbol{F}}(z^{-1}) = \boldsymbol{F}(z^{-1}) + \Delta\boldsymbol{F}(z^{-1}) \tag{19}$$

The computation of residuals may be based only on the model (the true plant is not known), so Eq. 10 becomes:

$$\boldsymbol{r}(t) = \hat{\boldsymbol{F}}(z^{-1})\tilde{\boldsymbol{q}}(t) = \boldsymbol{F}(z^{-1})\tilde{\boldsymbol{q}}(t) + \Delta\boldsymbol{F}(z^{-1})\tilde{\boldsymbol{q}}(t) \tag{20}$$

This changes the internal form (Eq. 11) of the residual equation to

$$\boldsymbol{r}(t) = \hat{\boldsymbol{F}}(z^{-1})\delta\boldsymbol{q}(t) - \boldsymbol{E}(z^{-1})\boldsymbol{w}(t) + \hat{\boldsymbol{F}}(z^{-1})\Delta\boldsymbol{q}(t) - \boldsymbol{G}_D(z^{-1})\boldsymbol{u}_D(t) + \Delta\boldsymbol{F}(z^{-1})\boldsymbol{q}(t) \tag{21}$$

That is, the coefficients associated with the measurement faults and noise are those of the model (and not of the true plant), and there is an additional modeling error term. The "coefficient" of the modeling error is the (true) variable vector $\boldsymbol{q}(t)$ indicating that the effect of the modeling errors on the residuals depends on the operating point.

By robustness of a parity equation or of an entire model we mean its insensitivity to modeling errors. Such robustness may be defined with the worst case error of all the model parameters in mind (Chow and Willsky, 1984) or with the assumption that there is a discrete set of uncertain models (Lou et al., 1986). We find it more practical to work with "partial robustness" with respect to certain underlying parameters. Such underlying parameters may be selected ones of the state space model or those of some physical/chemical description of the plant.

Partial robustness relates a single residual to a single underlying parameter. A direct measure of such robustness may be defined as the residual caused by the selected modeling error, if there are no failures and the instantaneous noise is zero, divided by the test threshold (Luo, 1989). Alternatively, the limit value of the parameter error that triggers the statistical test on the residual, with no failure or noise contributing, may be used as the measure of partial robustness (Gertler, 1986). In either case, the robustness measure depends on the test thresholds and on the operating point. For dynamic systems, the most recent $n + 1$ values of the plant variables need to be specified (n is the order of the system); usually it suffices to assume that they are constant.

Denote the vector of uncertain underlying parameters by θ. Then each element $\Delta f_{ij}(z^{-1})$ of the matrix $\Delta\boldsymbol{F}(z^{-1})$ is a function of $\Delta\theta$, the modeling error associated with θ. Thus a residual $r_i(t)$, obtained from Eq. 21 with all faults and noises zero, can be described as

$$r_i(\Delta\theta, \boldsymbol{q}) = \sum_j \Delta f_{ij}(\Delta\theta, z^{-1})q_j(t) \tag{22}$$

Introduce

$$\Delta\theta_l = [0 \ldots 0\ \Delta\theta_l\ 0 \ldots 0]^T \tag{23}$$

to describe the situation when all but one (the lth) of the underlying parameters are accurate. Then the direct measure of partial robustness with respect to this parameter is obtained as

$$\mu_i(\Delta\theta_l, \boldsymbol{q}) = \frac{|r_i(\Delta\theta_l, \boldsymbol{q})|}{\eta_i} \tag{24}$$

where η_i is the test threshold for $r_i(t)$. Finally, the limit value of $\Delta\theta_l$ that triggers the test on $r_i(t)$ is obtained as

$$\tau_{il}(\boldsymbol{q}) = \underset{\Delta\theta_l}{\text{sol}}\{|r_i(\Delta\theta_l, \boldsymbol{q})| = \eta_i\} \tag{25}$$

where sol $\{\ldots\}$ indicates the solution of the equation in the brackets. The solution may be obtained analytically or, if the relationship is strongly nonlinear, numerically. Note that, if the residual depends linearly on the parameter uncertainty, then the two measures are equivalent.

For a model consisting of several parity equations and with a number of underlying parameters selected, the robustness situation is described by a matrix of partial robustness measures. A model is as robust relative to a selected underlying parameter as its least robust equation.

Isolability Conditions

As we pointed out in the Introduction section, our primary concern in this work is the generation of models that guarantee failure isolability and are (sub)optimal for modeling error robustness. The concept and conditions of isolability will be outlined in this section; for more on the subject, refer to some earlier publications (Gertler and Singer, 1985; Gertler, 1988).

The residuals are subjected to statistical testing individually, each test yielding a boolean decision. Testing a set of residuals results in a boolean signature vector that provides a basis for the failure decision. In such a decision-making framework, the isolation of failures may be significantly enhanced if each residual is orthogonal to certain failures. Orthogonality means that the residual is not influenced by the given failure: that is, there is a zero coefficient at the concerned position of the system matrix. For example, each residual in Eq. 11 is orthogonal to all but one output sensor fault since the $\boldsymbol{H}(z^{-1})$ matrix is diagonal.

As can be seen, isolability is related closely to the structure of

the parity equations. This structure may be characterized best by the system's incidence matrices. An incidence matrix is a boolean representation of a model matrix, any zero model parameter represented by zero and any nonzero parameter by 1. Each column of the incidence matrix is associated with an input or output variable, that is, with a potential failure (cf. Eqs. 10 and 11). Ideally, the boolean signature in response to a single failure is identical with the respective column of the incidence matrix.

For any single failure to be detectable, the respective column of the incidence matrix must not be all zero. If there is no noise or the thresholds of the tests are set relatively low, then two single failures are isolable from each other provided their columns in the incidence matrix are different (low threshold or deterministic isolability).

Usually, the test thresholds are set high to reduce the occurrence of false alarms. In this situation, a fault of marginal size may trigger the test on some of the residuals while not triggering it on some others. To avoid false failure decisions when this happens, a special "high-threshold" ("statistically") isolabile model structure is necessary. High-threshold isolability is achieved most easily by "column canonical" incidence matrices, ones that have the same number of zeroes in each column, each in a different configuration.

Of the following three incidence matrices, Φ_1 signifies a system that is not isolable in either sense, Φ_2 one that is deterministically isolable but statistically not, and Φ_3 one that is isolable in both senses. The example also illustrates that isolability can be achieved by introducing appropriately-structured new parity equations (as substitutes or additions) to the original set.

$$\Phi_1 = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix} \quad \Phi_2 = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix} \quad \Phi_3 = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

The parity equations shown in Eqs. 10 and 11 are column canonical, that is, statistically (high-threshold) isolabile for the outputs but not for the inputs (in fact, they are not even low-threshold isolabile for the inputs). To achieve complete column canonical structures, a different set of parity equations needs to be generated. This will be discussed in a forthcoming section.

Generating Orthogonal Residuals

As we pointed out earlier, the idea of isolability hinges on the orthogonality of residuals to certain system variables and the associated faults. The residuals obtained in Eqs. 10 and 11 are orthogonal to all but one of the outputs (each one containing a different output), but in general they are not orthogonal to any input. To attain orthogonality to selected inputs, additional parity equations are necessary.

The generation of residuals with specific orthogonality properties may be posed as a transformation applied to the "primary" input-output models (Eqs. 10 and 11). This approach was first outlined in Gertler and Singer (1985) and later explored in more detail in Gertler and Singer (1990). The method is conceptually straightforward but computationally somewhat complex (among other things, it requires the finding of common factors among polynomials).

In the present paper, another method for generating orthogonal residuals will be discussed. It is based on a general procedure by Chow and Willsky (1984). The structural constraints by the isolability requirements will be superimposed on the original Chow-Willsky procedure. This approach is conceptually indirect but computationally simple. The resulting parity equations are the same as the ones obtained by model transformation.

Consider the state equations (Eq. 1) of the plant. The second equation describes the output $y(t)$ as

$$y(t) = Cx(t) + Du(t) \tag{26}$$

Now $y(t+1)$ may be obtained by substitution as

$$y(t+1) = CAx(t) + CBu(t) + Du(t) \tag{27}$$

Similarly $y(t+s)$ for any $s > 0$ is obtained as

$$y(t+s) = CA^s x(t) + CA^{s-1}Bu(t) + \cdots + CBu(t+s-1) + Du(t+s) \tag{28}$$

Collecting the equations for $s = 0 \ldots n$ (and shifting the time variable by n) yields the following scheme:

$$\begin{bmatrix} y(t-n) \\ y(t-n+1) \\ y(t-n+2) \\ \vdots \\ y(t) \end{bmatrix} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^n \end{bmatrix} x(t-n) + \begin{bmatrix} D & & & \\ CB & D & & \\ CAB & CB & & \\ \vdots & & & \\ CA^{n-1}B & CA^{n-2}B \cdots CB & D & \end{bmatrix} \begin{bmatrix} u(t-n) \\ u(t-n+1) \\ u(t-n+2) \\ \vdots \\ u(t) \end{bmatrix} \tag{29}$$

or in compact form

$$Y(t) = Qx(t-n) + RU(t) \tag{30}$$

For a system with m outputs and k inputs, vector $Y(t)$ is $(n+1) \times m$ long and vector $U(t)$ is $(n+1) \times k$ long. Matrix Q has $(n+1) \times m$ rows and n columns while matrix R has $(n+1) \times m$ rows and $(n+1) \times k$ columns.

Premultiplying Eq. 30 with a row vector ω^T, $(n+1) \times m$ long, yields a scalar equation:

$$\omega^T Y(t) = \omega^T Qx(t-n) + \omega^T RU(t) \tag{31}$$

In general, this equation will contain a mix of input, output and state variables. It will qualify as a parity equation only if the state variables disappear. This requires that

$$\omega^T Q = 0 \tag{32}$$

That is, the $(n+1) \times m$ elements in ω^T have to satisfy a set of n

homogeneous equations. If the system is observable, these n equations are independent. (Matrix $\boldsymbol{Q}$ may be recognized as the observability matrix, with one row, $\boldsymbol{CA}^n$ added. If the system is observable, $\boldsymbol{Q}$ has rank n: that is, the n columns are independent.) Therefore, $(n+1) \times m - n$ elements may be chosen freely. However, to obtain a nontrivial solution, at least one of the free elements must be nonzero. This leaves $(n+1) \times (m-1)$ potential zero elements.

To make a parity equation orthogonal to an output $y_i(t)$, all occurrences of y_i have to disappear from the equation: that is, $y_i(t), y_i(t-1), \ldots, y_i(t-n)$. This requires $n+1$ zeroes in ω^T, appropriately positioned. Thus a parity equation may be made orthogonal to up to $m-1$ outputs. In fact, choosing one output at a time to stay, this technique regenerates the input-output equations (Eq. 5). In other words, it produces the "primary" parity equations.

Orthogonality to an input $u_j(t)$ requires zeroes in the vector

$$\zeta^T = \omega^T \boldsymbol{R} \tag{33}$$

To eliminate all occurrences of u_j takes $n+1$ appropriately-positioned zeroes: that is, $n+1$ homogeneous equations on the elements of ω^T. This adds to the original n homogeneous equations and reduces the number of free elements. The total number of variables, inputs and ouputs, to which a parity equation may be made orthogonal, is $m-1$.

Orthogonalization to inputs implies manipulations with columns of matrix $\boldsymbol{R}$. The good behavior of these columns in general is not guaranteed. Irregularities in matrix $\boldsymbol{R}$ (zero column or linear dependence) may lead to the abortion of the procedure or to orthogonality to more variables than intended. In either case, the desired parity equation is not attainable.

If each parity equation is orthogonal to exactly $m-1$ variables, the set of equations is called "row canonical," meaning that each equation has the same number of zeroes, each in a different configuration. A row-canonical structure has no direct bearing on deterministic or statistical isolability but may be advantageous in generating models that possess those properties. This will be explored in the next section.

Note that the procedure described above makes no distinction between accessible (measured or controlled) inputs and additive plant disturbances. This is desirable since orthogonalization is to apply to the latter as well. For the computation of residuals, however, the disturbance variables have to be dropped from the resulting equations.

Generating Diagnostic Models

As described earlier, a diagnostic model is a consistent set of parity equations. For isolability, the model should be in the appropriate column canonical structure. There is a combinatorial multitude of models satisfying this requirement. The selection can be made on the basis of a combination of performance measures, including

- Sensitivity condition
- Maximum (or average) sensitivity number
- Partial robustness measures

The sensitivity condition and partial robustness measures characterize single parity equations. When the equations form a diagnostic model, the poorest performer determines the model performance. The maximum (or average) sensitivity number, by specifying the necessary threshold adjustment for the design of residual filters, influences the detection delays. In a diagnostic model, the isolation decision is held up until the slowest equation detects a threshold passing (meanwhile, the failure signature is undefined). Therefore, equations with smaller and/or uniform delays should be preferred in the design.

Some of the above measures may be used to reduce the set of useful equations by setting limits on the acceptable performance while others may serve as a basis for optimization. The larger the set of parity equations, the more flexibility the designer has in selecting the diagnostic model.

A reasonable design strategy is to set limits on the sensitivity condition (and perhaps the maximum sensitivity number) and eliminate the outlying equations, then perform optimization, constrained to satisfy the structural requirements, over the reduced set with respect to different partial robustness measures. This leads to a family of diagnostic models, each column canonical and within the prescribed sensitivity limits, and each one optimally robust within those constraints relative to a selected parameter uncertainty.

As the first step of model design, the complete set of parity equations is to be generated. To simplify the design problem, only row-canonical equation structures are considered. For m outputs and k inputs, there are

$$P = \binom{m+k}{m-1} \tag{34}$$

such structures. However, as mentioned before, zeroes and linear dependences in the system matrices may render some of the parity equations unattainable and thus irrelevant to model generation. Even so, apart from trivially small systems, the attainable set may be quite extensive (for example, for $m=4$ and $k=6$, $P=120$; but for $m=6$ and $k=8$, $P=2184$).

Subsequently, the selected performance measures are computed for each equation. If limits are set for some of the measures, then the set is reduced by eliminating the outlying equations. The remaining set is then ordered according to the measure chosen as a basis for optimization (if there are a number of such measures, the same number of ordered sets is formed).

Optimization for a selected performance measure within the framework of isolability requires a search over the available (reduced) set of equations under the constraint of column-canonical model structure. Two approaches will be discussed below. One implies a global search over the set; this guarantees the optimality of the model, but its computational cost may be prohibitive. The other approach utilizes a local search procedure; this is computationally attractive, but may lead to a local optimum.

When forming column-canonical models of row-canonical equations, the number p of equations in the model and the number s of zeroes per column have to satisfy the following relationships:

$$p(m-1) = s(m+k)$$
$$\binom{p}{s} \geq m+k \tag{35}$$

Here the equation simply expresses that the total number of zeroes in the rows is identical with that in the columns while the inequality guarantees that a sufficient number of different col-

umns may be created. While there are many possible solutions, the one yielding the simplest diagnostic model is to be sought.

The numerical coefficients of the parity equations are needed for the performance computations, but not for the subsequent search. The memory requirements of the algorithm may be reduced significantly, if the coefficients are not stored, only the equation structure and the performance measures.

Global search procedure

The global search procedure operates on the ordered set of equations starting with the best performers. The first p equations are selected and checked for column structure. If this subset is column canonical, it represents the best model. Otherwise, the investigated subset is expanded to include the first $p + 1$ equations. All the possible p-tuples are formed and checked for column structure; if any of them qualifies, it is declared as the best model. Otherwise, the subset is expanded to $p + 2$, $p + 3$, etc. equations, until the first column-canonical p-tuple is found.

Some savings in the search results from the observation that the p-tuples to be checked always contain the last equation (since all possibilities without the last equation have been exhausted before). Thus in the stage when the investigated subset contains $p + i$ equations, the number of p-tuples to be checked is $\binom{p+i-1}{p-1}$ while the total number of p-tuples checked up to and including this ith stage is $\binom{p+i}{p}$. This number may grow extremely large if a column-canonical model is not found early in the search. Though usually there is a combinatorial number of suitable structures, since the equations are ordered according to performance measures (that depend on numerical coefficients), their distribution over the set of equations does not follow any predictable rule. Therefore there is no guarantee that the first suitable model would be found sufficiently early. (For example, with $m = 4$ and $k = 6$, covering one quarter of the equation set would require checking of about 3×10^7 model structures, that is quite feasible to handle. With $m = 6$ and $k = 8$, going over 2% of the list implies only about 2×10^6 models but covering 10% imples 4×10^{21}, that is obviously not tractable by the present-day computer technology.)

Local search procedure

The local search procedure starts from an arbitrary column-canonical model and uses a series of systematic row replacements to improve its performance while maintaining the column-canonical structure.

The procedure utilizes the concept of structural distance between row-canonical equations. Interchanging a 0 and a 1 in an equation structure with $m - 1$ zeroes results in another structure with the same number of zeroes. The minimum number of such 0–1 interchanges needed to bring one structure into another is called the structural distance between the two equations. For example, the structural distance between 111100 and 111010 is one, while between 111100 and 001111 it is two. The distance can be conveniently obtained by halving the number of 1's in the exclusive OR of the two boolean patterns. The maximum possible distance in a system with k inputs and m outputs is $min\ [k + 1, m - 1]$. For a set of equations, a distance matrix may be set up, at least conceptually, showing the structural distances between all possible pairs.

Given any column-canonical initial model, the equation with the poorest performance measure is selected for replacement. In the basic form of the algorithm, only equations with distance 1 from the original equation will be considered as substitutes; there are $(k + 1) \times (m - 1)$ such potential substitutes for any equation. The search starts with the distance 1 substitute that has the best performance. The candidate substitute equation disqualifies if it is already a member of the model or if its performance is worse than that of the equation to be replaced.

Replacing a single equation would spoil the column-canonical structure; a complementary replacement is necessary to restore it. There may be several equations in the initial model that, if appropriately replaced, restore the structure. For each of these, a distance 1 substitute is sought, with the best possible performance, and the one with the best overall performance is then chosen. A candidate for complementary substitute disqualifies if it is already in the model or if its performance is worse than that of the equation originally replaced (i.e., of the worst equation in the initial model). If no suitable complementary substitute is found, the original replacement has to be modified, moving to the substitute equation with the next best performance.

Once the equation with the poorest performance has been successfully replaced, the procedure is repeated for the worst equation in the resulting model and so on. The search stops if no distance 1 replacement (with distance 1 complementary replacement) can be found for the worst equation in the latest model.

Since the search is limited to distance 1 replacements, it may easily stop in a local optimum. Convergence toward the global optimum can be improved by allowing distance 2, 3, etc. replacements, preferably once all the distance 1 replacements have been exhausted. (In fact, after a higher distance replacement, a series of distance 1 steps may become possible again.) This, however, requires a significantly more complex algorithm; especially the complementary replacements become rather complicated.

The column-canonical initial model needed to start the search may be generated according to some arbitrary symmetrical pattern. If it turns out to contain equations that are unattainable for the given system, other symmetrical patterns may be attempted. Alternatively, the unattainable equations may be assigned some very poor (arbitrary) performance measures, thus making them the first candidates for replacement.

Further design considerations

The procedure to design isolabile diagnostic models with implied sensitivity and robustness considerations is rather complex and cannot be completely automated. At some points, the human designer has to interact with the design program and make certain decisions. These require a good understanding of the meaning of different design parameters and of the implications of modifying them.

One of the decisions is related to the truncation values of the different performance measures. If truncation does not leave a sufficiently large equation pool for model selection, then the designer may want to relax the truncation values or otherwise modify the specification (e.g., by relaxing some nominal failure sizes).

A deeper issue concerns the trade-off between sensitivity and detection/isolation delay and imples such considerations as the selection of the sampling interval and residual filtering. With the discretized model and the noise situation given, the residual

variances follow; these, with a selected false alarm ratio, determine the test thresholds. However, the choice of the sampling interval influences the parameters of the discretized model and has a significant effect on the thresholds and sensitivities. Increasing the sampling interval, in general, leads to smaller residual variances and better sensitivity numbers (while leaving the sensitivity conditions uneffected) at the expense of slower reaction to failures. Filtering of the residuals is another way of improving failure sensitivity, again at the expense of delayed detection/isolation. While a longer sampling interval affects the entire model, residual filters are designed separately for each parity equation, tailored to the conditions of the particular equation. However, filtering does not distinguish between failure and modeling error, unless they are in different frequency ranges; improvements in failure sensitivity are usually accompanied by deterioration of modeling error robustness.

The partial robustness design concept described in this paper has, of course, certain limitations. Any diagnostic model is optimized only for some selected modeling uncertainty in a selected operating point. The behavior of models under different circumstances may be analyzed and the thresholds increased, if necessary, to improve modeling error robustness; this, however, would adversely effect failure sensitivity. An important implication of isolable model structures is that they provide a certain kind of inherent robustness against modeling error effects as well, in that it is extremely unlikely that modeling errors would lead to any valid failure signature and thus be misinterpreted as a particular failure. With multiple parallel diagnostic models, if designed with realistic modeling error situations in mind, there is a good chance that one or more of the models is also sufficiently shielded from modeling error effects and thus provides a clear failure indication at any given time. The combination of incomplete and/or contradictory evidence from parallel models has been the subject of a related research effort (Gertler and Anderson, 1989).

As pointed out earlier, the global search for the best diagnostic model (under a selected performance measure) may easily become computationally intractable because of the combinatorial nature of the problem. The local search procedure introduced in this paper is computationally efficient and places practically no limitation on the size of systems that can be dealt with, but it does not guarantee the optimality of the solution. As in many other engineering problems, one may have to settle here for acceptability (instead of optimality); if the designer is not satisifed with the performance measures of a locally optimal model, a new local search may be initiated from a different arbitrary initial set of equations. The local search procedure may be improved at the expense of increased computational complexity, by allowing higher distance replacements or by including some annealing technique.

Design Example

The design procedure will be demonstrated on a distillation column model (Takamatsu et al., 1979). The column has nine plates. The flow rate of the reflux and of the vapor stream are controlled based on the liquid composition at the second and ninth plate. Variations in the feed flow rate and composition are unmeasured disturbances.

The plant is described by a linear state-space model, obtained by linearization around a nominal steady-state operating point. The model has the following properties:

- *Eleven state variables*
  - distillate composition
  - nine plate compositions
  - bottom product composition
- *Two controlled inputs*
  - $u_1(t)$: reflux flow rate
  - $u_2(t)$: vapor stream flow rate
- *Two disturbance inputs*
  - $u_3(t)$: feed stream composition
  - $u_4(t)$: feed stream flow rate
- *Two outputs*
  - $y_1(t)$: distillate composition
  - $y_2(t)$: bottom product composition

All variables in the linear model are changes relative to the nominal steady-state operating point. The main characteristics of the nominal operating point are summarized in the Appendix.

The Eigenvalues of the continuous-time state-space model are all real and range from 0.02 1/min to 3.3 1/min, corresponding to time constants between 50 min and 0.3 min. The model was discretized using a 8-min sampling interval. Shorter sampling intervals would lead to poor sensitivity performance of the parity equations; much shorter ones also lead to numerical difficulties in the discrete representation. The resulting values for the discrete system matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ (cf. Eq. 1) are also given in the Appendix. Of the other matrices, $\boldsymbol{D} = \boldsymbol{O}$ and $\boldsymbol{L} = \boldsymbol{I}$.

With $k = 4$ inputs and $m = 2$ outputs, there is a total of $\binom{6}{1} = 6$ row-canonical equations, that is a trivially simple system. Two of the plate compositions are sensed and used for control; considering these as additional outputs would lead to a system with $k = 4$ inputs and $m = 4$ outputs, resulting in a total of $\binom{8}{3} = 56$ row-canonical equations. Such a selection of parity equations may still be insufficient to allow a full-blown design for a combination of sensitivity and robustness measures. To demonstrate the potentials of the methodology, let us include two more state variables as outputs. Thus there is a total of six outputs:

$y_1(t)$: composition of the distillate
$y_2(t)$: liquid composition on the 2nd plate
$y_3(t)$: liquid composition on the 4th plate
$y_4(t)$: liquid composition on the 7th plate
$y_5(t)$: liquid composition on the 9th plate
$y_6(t)$: composition of the bottom product

The corresponding $\boldsymbol{C}$ matrix is shown in the Appendix. Now with $k = 4$ inputs and $m = 6$ outputs, the full row-canonical set contains $\binom{10}{5} = 252$ parity equations. The smallest column-canonical model that may be formed of these equations consists of $p = 6$ rows (equations) with $s = 3$ zeroes per column.

The system, sensor and actuator noises are all assumed to be normal, white (uncorrelated in time) with zero mean and independent of one another. Their variances were established (guessed) on the basis of the character and nominal value of the variables. The nominal detectable failure sizes were specified using similar considerations. The noise variances and nominal failures are shown in Table 1. Of course, the validity of the selected values may be argued: they might be placed on a more solid ground utilizing detailed process knowledge and noise measurements.

The test thresholds are set at four times the standard deviation of the respective residual.

As uncertain underlying parameters, the two columns of the $\boldsymbol{B}$ matrix that belong to the control inputs have been chosen.

Table 1. Noise Variances and Nominal Failures*

	Noise Variance	Nominal Failure
State Variables	1×10^{-8}	—
Inputs		
u_1	6.76×10^{-4}	0.75
u_2	11.70×10^{-4}	0.75
u_3	—	0.02
u_4	—	0.50
Outputs		
y_1	4×10^{-6}	0.025
y_2	4×10^{-6}	0.025
y_3	3×10^{-6}	0.025
y_4	2×10^{-6}	0.010
y_5	1×10^{-6}	0.010
y_6	1×10^{-6}	0.010

*All units conform with Appendix 2.

This choice is supported by the common understanding that under linearization the model gains may be at error if we depart from the nominal operating point. One full column at a time is assumed to have modeling error, the same 10% for each element. Robustness is investigated at the temporary operating point characterized by

$$u_1(t) = 1.30 \quad u_2(t) = 1.74$$

that corresponds to a 10% departure from the nominal steady-state operating point.

The design program generates the full row-canonical set of 252 parity equations. Also, for each equation, the following measures are computed:

- Test threshold η_i
- Sensitity condition κ_i
- Maximum sensitivity number λ_{imax}
- Robustness measure μ_{i1} relative to the first column of $\boldsymbol{B}$
- Robustness measure μ_{i2} relative to the second column of $\boldsymbol{B}$

The sensitivity condition κ_i for the 252 equations ranges from 1.35 to 1.4×10^8. The extremely high sensitivity conditions apparently signify unattainable equations; excess zeroes are checked numerically in the transformation procedure and computational inaccuracies may cause some zero coefficients seem nonzero. The maximum sensitivity number λ_{imax} ranges from 1.23 to 8×10^8; the extremely high numbers appear in the equations that have very high condition numbers as well. Choosing 15 as the maximum acceptable sensitivity condition reduces the number of useful equations to 97, with a λ_{imax} range of 1.23 to 115. Eliminating also the equations for which $\lambda_{imax} > 25$ leaves 72 equations.

Table 3. Locally-Optimal Model for Sensitivity Condition κ_i $\lambda_{imax} < 25$

Structure					
Outputs	Inputs	κ_i	λ_{imax}	μ_{Fi1}	μ_{Fi2}
0 1 1 1 0 0	1 0 0 1	1.819	1.572	0.175	0
0 0 0 1 1 0	1 1 1 0	1.841	9.352	0.169	0.171
0 1 0 0 0 1	1 1 0 1	3.126	1.579	0.098	0.135
1 0 0 0 1 1	0 1 0 1	5.112	4.958	0	0.103
1 0 1 1 0 1	0 0 1 0	8.055	2.914	0	0
1 1 1 0 1 0	0 0 1 0	10.650	5.504	0	0

A sensitivity number exceeding 1 indicates that the concerned fault, occurring at its nominal size, would not trigger the test on the concerned residual. In order to reduce the thresholds the residuals are filtered. The new thresholds are selected as

$$\eta_{Fi} = \eta_i \frac{\sqrt{\kappa_i}}{\lambda_{imax}} = \frac{\eta_i}{\sqrt{\lambda_{imin} \cdot \lambda_{imax}}}$$

This way, the minimal and maximal responses to nominal size failures are symmetrical relative to the threshold in any equation.

Residual filtering does not change the sensitivity condition, but it does modify the robustness measures. Therefore, before optimization for robustness, the robustness measures μ_{Fi1}, μ_{Fi2} for the filtered residuals were computed, only for the 72 equations with $\kappa_i < 15$ and $\lambda_{imax} < 25$. The filtered robustness measures range from 1.6×10^{-8} to 0.538 relative to uncertainties in the first column of the $\boldsymbol{B}$ matrix and from 1.3×10^{-5} to 0.541 relative to those in the second column. (The reader should be reminded here that these are inverse robustness measures: the smaller the number, the better the performance.) The extremely small, but nonzero, values are due to computational inaccuracies; for any equation that does not contain the first (second) input variable, the robustness measure relative to the first (second) column should be zero. The truly nonzero values span from 0.055 to 0.538 and from 0.065 to 0.541, respectively.

Both a global and a local search was performed, the latter initiating from an arbitrary column canonical model, over the 72 equations with the sensitivity condition κ_i as performance index. The results are shown in Table 2 and 3. As can be seen, the sensitivity condition of the worst equation is almost twice as large in (this particular) local optimum than in the globally-optimal model. However, the global optimum took about an hour of

Table 2. Globally-Optimal Model for Sensitivity Condition κ_i $\lambda_{imax} < 25$

Structure					
Outputs	Inputs	κ_i	λ_{imax}	μ_{Fi1}	μ_{Fi2}
0 0 0 1 1 0	1 1 1 0	1.841	9.352	0.169	0.171
0 0 1 0 0 1	1 0 1 1	3.285	9.139	0.139	0
0 1 0 1 0 0	1 1 0 1	4.119	5.504	0.265	0.114
1 1 1 1 1 0	0 0 0 0	4.550	5.439	0	0
1 0 1 0 1 1	0 0 1 0	5.950	6.338	0	0
1 1 0 0 0 1	0 1 0 1	5.967	1.843	0	0.095

Table 4. Locally-Optimal Model for Robustness Measure μ_{Fi1} $\lambda_{imax} < 25$

Structure					
Outputs	Inputs	κ_i	λ_{imax}	μ_{Fi1}	μ_{Fi2}
1 1 1 1 1 0	0 0 0 0	4.550	5.439	0	0
1 0 0 1 1 0	0 1 0 1	3.837	7.058	0	0.179
0 1 1 0 1 1	0 0 1 0	6.526	2.154	0	0
1 0 1 0 0 0	1 0 1 1	13.200	23.190	0.055	0
0 1 0 0 0 1	1 1 0 1	3.126	1.579	0.098	0.135
0 0 0 1 0 1	1 1 1 0	2.142	13.620	0.157	0.159

Table 5. Evolution of Robustness Measure μ_{Fi1} over Local Search

Eq.	Initial	1st Step	2nd Step	3rd Step	4th Step
1	0	0	0	0	0
2	0	0	0	0	0
3	0	0	0	0	0
4	0.139	0.139	0.055	0.055	0.055
5	0.169	0.169	0.169	0.098	0.098
6	0.265	0.199	0.172	0.169	0.157

mainframe computer time to find while the local optimum was found in about a minute.

A local search was performed with the first robustness measure μ_{Fi1} as performance index, starting from the model globally optimal for the sensitivity condition. The resulting model is shown in Table 4. The evolution of the robustness measures over the four steps of the local search procedure is demonstrated in Table 5.

Local search was also attempted, using the same initial model as above, for the second robustness measure μ_{Fi2} and the maximum sensitivity number λ_{imax} but no improvement could be achieved. With a different (arbitrary) initial model, the local search procedure yielded an improved model for λ_{imax} as shown in Table 6.

As the final step of the design procedure, the equation coefficients for the optimal diagnostic models are computed. The coefficients are polynomials in the shift operator. The maximum degree of the polynomials depends on the order of the system and the number of outputs in the particular equation. For the 11th-order system at hand with one to five outputs in the equations this maximum degree ranges from 7 to 11. The polynomials obtained by computation are generally of lower degree than these maxima.

Conclusion

A methodology has been introduced to design robust diagnostic models for complex systems described by linear discrete dynamic equations. The models consist of sets of parity equations that are rearranged and transformed versions of the input-output equations.

The design methodology is built around the concept of isolability and a set of equation performance measures. The isolability concept leads to a structural constraint on the dignostic model, expressed in terms of the position of zeroes in the parity equations. The performance measures characterize the sensitivity of each parity equation relative to different failures and the robustness of the same with respect to uncertainties in an underlying model.

Table 6. Locally-Optimal Model for Max. Sensitivity Number λ_{imax}; $\lambda_{imax} < 25$

Structure					
Outputs	Inputs	κ_i	λ_{imax}	μ_{Fi1}	μ_{Fi2}
0 1 1 1 0 0	1 0 0 1	1.819	1.572	0.175	0
0 0 1 0 1 0	1 1 0 1	2.550	1.990	0.109	0.325
1 1 1 0 1 0	0 0 1 0	10.650	5.504	0	0
1 0 0 1 0 1	0 1 0 1	5.746	6.377	0	0.218
0 1 0 0 0 1	1 1 1 0	6.275	6.557	0.434	0.437
1 0 0 1 1 1	0 0 1 0	7.691	6.728	0	0

In general, there is a large number of parity equations that can serve as elements for a diagnostic model and a combinatorial multitude of models that satisfy the structural requirements. By placing limits on some of the performance measures, the set of useful equations may be reduced considerably. Some other performance measures may be chosen as optimization criteria in a model search over the reduced set under the structural constraint. With an increasing system size, the design flexibility grows significantly, together with an approximately linear growth in the computational cost.

The potentials of the design methodology and the program package implementing it have been demonstrated on the linearized, 11th-order model of a distillation column with four inputs and six outputs.

Literature Cited

Almasy, G. A., and T. Sztano, "Checking and Correction of Measurements on the Basis of Linear System Model," *Problems of Control and Information Theory*, **4**, 57 (1975).

Ben-Haim, Y., "An Algorithm for Failure Location in a Complex Network," *Nucl. Sci. and Eng.*, **75**, 191 (1980).

———, "Malfunction Isolation in Linear Stochastic Systems: Application to Nuclear Power Plants," *Nucl. Sci. and Eng.*, **85**, 156, (1983).

Chow, E. Y., and A. S. Willsky, "Analytical Redundancy and the Design of Robust Failure Detection Systems," *IEEE Trans. on Autom. Cont.*, **AC-29**, 603 (1984).

Crowe, C. M., "Recursive Identification of Gross Errors in Linear Data Reconcillation," *AIChE J.*, **34**, 541 (1988).

Gertler, J., Gy. Varga, and F. Sipos, "Computational Experience with the Constrained Minimum Variance Input-Output Estimator," *Automat.*, **16**, 423 (1980).

Gertler, J., and D. Singer, "Augmented Models for Statistical Fault Isolation in Complex Dynamic Systems," *Proc. Amer. Control Conf.*, 317, Boston (1985).

Gertler, J., "A Numerical-Structural Approach to Failure Detection and Isolation in Complex Plants," *Proc. IEEE Conf. on Decision and Control*, 1576, Athens, Greece (1986).

———, "A Survey of Model-Based Failure Detection and Isolation in Complex Plants," *IEEE Control Systems Mag.*, 3, (Dec., 1988).

Gertler, J., and K. Anderson, "An Evidential Reasoning Extension to Model Based Failure Detection and Isolation," *IEEE Symp. on Intelligent Control Systems*, Albany, NY (1989).

Gertler, J., and D. Singer, "A New Structural Framework for Parity Equation Based Failure Detection and Isolation," *Automat.*, to be published (Mar., 1990).

Iordache, C., R. S. H. Mah, and A. C. Tamhane, "Peformance Studies of the Masurement Test for Detection of Gross Errors in Process Data," *AIChE J.*, **31**, 1187 (1985).

Kramer, M. A., "Malfunction Diagnosis Using Quantitative Models with Non-Boolean Reasoning in Expert Systems," *AIChE J.*, **33**, 130 (1987).

Lou, X. C., A. S. Willsky, and G. C. Verghese, "Optimally Robust Redundancy Relations for Failure Detection in Uncertain Systems," *Automat.*, **22**, 333 (1986).

Luo, Q., "Parity Equation Approach to Failure Detection and Isolation in Dynamic Systems," MS Thesis, George Mason Univ., Fairfax, VA (1989).

Mah, R. S. H., G. M. Stanley, and D. M. Downing, "Reconcillation and Rectification of Process Flow and Inventory Data," *Ind. and Eng. Chemistry, Proc. Des. Dev.*, **15**, 175 (1976).

Potter, J. E., and M. C. Suman, "Thresholdless Redundancy Management with Arrays of Skewed Instruments," *AGARDograph*, NATO, **224**, 15-1 (1977).

Romagnoli, J.A., and G. Stephanopoulos, "Rectification of Process Measurement Data in the Presence of Gross Errors," *Chem. Eng. Sci.*, **36**, 1849 (1981).

Takamatsu, T., I. Hashimoto, and Y. Nakai, "A Geometric Approach

to Multivariable Control System Design of a Distillation Column" *Automat.*, **15**, 387 (1979).
Vaclavek, V., "Gross Systematic Errors or Biases in the Balance Calculations," *Papers of the Prague Institute of Technology*, Prague, Czechoslovakia (1974).

Appendix: Residual Variance under Filtering

Consider a parity equation driven only by noise (cf. Eq. 12)

$$r_i(t) = f_{i\cdot}(z^{-1})\delta\boldsymbol{q}(t) - e_{i\cdot}(z^{-1})\boldsymbol{w}(t) \qquad \text{(A1)}$$

where $\delta\boldsymbol{q}(t)$ has $m + k$ elements and $\boldsymbol{w}(t)$ has ν elements and

$$f_{i\cdot}(z^{-1}) = f_{0i\cdot} + f_{1i\cdot}z^{-1} + \cdots + f_{ni\cdot}z^{-n}$$
$$e_{i\cdot}(z^{-1}) = e_{0i\cdot} + e_{1i\cdot}z^{-1} + \cdots + e_{ni\cdot}z^{-n} \qquad \text{(A2)}$$

are the ith row of $\boldsymbol{F}(z^{-1})$ and $\boldsymbol{E}(z^{-1})$, respectively. The residual $r_i(t)$ is filtered by a second-order filter (cf. Eq. 15).

$$r_{Fi}(t) = \frac{1 + \gamma_1 + \gamma_2}{1 + \gamma_1 z^{-1} + \gamma_2 z^{-2}} r_i(t) \qquad \text{(A3)}$$

Table A1. Steady-State Operating Point

Flow Rate	
Feed	10.00 mol/min
Distillate	4.35 mol/min
Bottom Product	5.65 mol/min
Vapor Stream	17.40 mol/min
Reflux	13.05 mol/min
Liquid Composition	
Feed	0.54036
Condenser	0.90000
2nd Plate	0.72211
4th Plate	0.52804
7th Plate	0.32005
9th Plate	0.16772
Reboiler	0.10400

Assume that $\delta\boldsymbol{q}(t)$ and $\boldsymbol{w}(t)$ are white, independent of each other and, separately, have normal multivariate distributions with zero mean and known covariances $\boldsymbol{S}_q$ and $\boldsymbol{S}_w$, respectively. Then the variance of $r_{Fi}(t)$ may be computed the following way (Gertler et al., 1981):

$$\text{var}\{r_{Fi}(t)\} = \sum_{j=1}^{m+k}\sum_{l=1}^{m+k} \pi_{ijl}s_{qjl} + \sum_{j=1}^{\nu}\sum_{l=1}^{\nu} \omega_{ijl}s_{wjl} \qquad \text{(A4)}$$

Here s_{qjl} and s_{wjl} are the jlth elements of $\boldsymbol{S}_q$ and $\boldsymbol{S}_w$, respectively, and π_{ijl} and ω_{ijl} are elements of the matrices $\boldsymbol{\Pi}_i$ and $\boldsymbol{\Omega}_i$, computed as

$$\boldsymbol{\Pi}_i = (1 + \gamma_1 + \gamma_2)^2 \sum_{p=0}^{n}\sum_{g=p}^{n}\sum_{h=0}^{n} \phi_{0p}\psi_{(g-p)h}\boldsymbol{f}^T_{gi\cdot}\boldsymbol{f}_{hi\cdot} \qquad \text{(A5)}$$

$$\boldsymbol{\Omega}_i = (1 + \gamma_1 + \gamma_2)^2 \sum_{p=0}^{n}\sum_{g=p}^{n}\sum_{h=0}^{n} \phi_{0p}\psi_{(g-p)h}\boldsymbol{e}^T_{gi\cdot}\boldsymbol{e}_{hi\cdot} \qquad \text{(A6)}$$

where, further, ϕ_{0p} and $\psi_{(g-p)h}$ are elements of the matrices $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$ defined as

$$\boldsymbol{\Psi}^{-1} = \begin{bmatrix} 1 & 0 & 0 & & & & & & & 0 \\ \gamma_1 & 1 & 0 & 0 & & & & & & 1 \\ \gamma_2 & \gamma_1 & 1 & 0 & & & & & & 2 \\ 0 & \gamma_2 & \gamma_1 & 1 & \cdot & & & & & \cdot \\ 0 & 0 & \gamma_2 & \gamma_1 & \cdot & \cdot & & & & \cdot \\ & 0 & 0 & \gamma_2 & \cdot & \cdot & 0 & 0 & & \cdot \\ & & & & \cdot & \cdot & 1 & 0 & 0 & \cdot \\ & & & & & \cdot & \gamma_1 & 1 & 0 & \cdot \\ & & & & & & \gamma_2 & \gamma_1 & 1 & n \\ 0 & 1 & 2 & \cdot & \cdot & \cdot & \cdot & \cdot & n & \end{bmatrix}$$

Table A2. A Matrix

4.4963e-01	8.8870e-02	6.4123e-02	4.2937e-02	2.6130e-02	1.2237e-02
6.3217e-03	3.2826e-03	1.6940e-03	7.9448e-04	1.2263e-03	
4.4181e-01	1.1755e-01	1.1147e-01	9.4334e-02	6.9634e-02	3.9503e-02
2.4238e-02	1.4617e-02	8.5404e-03	4.4300e-03	8.6760e-03	
3.5481e-01	1.2407e-01	1.3976e-01	1.3372e-01	1.0878e-01	6.8529e-02
4.6153e-02	3.0182e-02	1.8879e-02	1.0390e-02	2.3703e-02	
2.3444e-01	1.0361e-01	1.3196e-01	1.3885e-01	1.2346e-01	8.7199e-02
6.5078e-02	4.6540e-02	3.1396e-02	1.8521e-02	5.0610e-02	
1.2495e-01	6.6982e-02	9.4014e-02	1.0813e-01	1.0607e-01	8.6128e-02
7.2308e-02	5.7060e-02	4.1805e-02	2.6731e-02	8.9454e-02	
5.1252e-02	3.3281e-02	5.1873e-02	6.6890e-02	7.5434e-02	7.4216e-02
7.1700e-02	6.3117e-02	5.0663e-02	3.5654e-02	1.4903e-01	
3.4000e-02	2.6222e-02	4.4861e-02	6.4105e-02	8.1324e-02	9.2072e-02
9.7105e-02	9.1320e-02	7.7731e-02	5.8652e-02	2.8496e-01	
2.0373e-02	1.8247e-02	3.3854e-02	5.2901e-02	7.4053e-02	9.3528e-02
1.0538e-01	1.0471e-01	9.4393e-02	7.6958e-02	4.3577e-01	
1.0758e-02	1.0910e-02	2.1669e-02	3.6519e-02	5.5521e-02	7.6822e-02
9.1788e-02	9.6593e-02	9.3233e-02	8.3604e-02	5.5560e-01	
4.5565e-03	5.1105e-03	1.0769e-02	1.9454e-02	3.2058e-02	4.8821e-02
6.2542e-02	7.1115e-02	7.5497e-02	7.6491e-02	5.9897e-01	
7.0824e-04	1.0079e-03	2.4740e-03	5.3533e-03	1.0803e-02	2.0550e-02
3.0599e-02	4.0551e-02	5.0524e-02	6.0316e-02	5.5732e-01	

Distillation column data in the design example are shown in Tables A1–A4 on pp. 1867–1868.

$$\Phi^{-1} = \Psi^{-1} + \begin{bmatrix} 0 & \gamma_1 & \gamma_2 & 0 & \cdot & \cdot \\ 0 & \gamma_2 & 0 & 0 & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \quad \text{(A7)}$$

If the covariance matrices S_q and S_w are diagonal, Eq. A4 simplifies to

$$\text{var}\{r_{Fi}(t)\} = \sum_{j=1}^{m+k} \pi_{ijj} \,\text{var}\{\delta q_j(t)\} + \sum_{j=1}^{\nu} \omega_{ijj} \,\text{var}\{w_j(t)\} \quad \text{(A8)}$$

with

$$\pi_{ijj} = (1 + \gamma_1 + \gamma_2)^2 \sum_{p=0}^{n} \sum_{g=p}^{n} \sum_{h=0}^{n} \phi_{0p}\psi_{(g-p)h} f_{gij} f_{hij} \quad \text{(A9)}$$

$$\omega_{ijj} = (1 + \gamma_1 + \gamma_2)^2 \sum_{p=0}^{n} \sum_{g=p}^{n} \sum_{h=0}^{n} \phi_{0p}\psi_{(g-p)h} e_{gij} e_{hij} \quad \text{(A10)}$$

where f_{gij} and e_{gij} are the jth element of $f_{gi\cdot}$ and $e_{gi\cdot}$, respectively.

Table A3. *B* Matrix

min/mol	min/mol	1	min/mol
4.8246e-03	−3.6506e-02	1.3907e-02	6.9512e-05
1.6355e-02	−1.2492e-02	6.9663e-02	4.2344e-04
2.4180e-02	−1.8772e-02	1.5828e-01	1.0761e-03
2.6420e-02	−2.1106e-02	2.7959e-01	2.1515e-03
2.3048e-02	−1.9381e-02	4.1293e-01	3.6167e-03
1.7560e-02	−1.6560e-02	5.8752e-01	5.8606e-03
2.0132e-02	−2.1573e-02	4.8018e-01	1.1221e-02
2.0736e-02	−2.3960e-02	3.5373e-01	1.4566e-02
1.8229e-02	−2.2040e-02	2.2581e-01	1.4488e-02
1.2688e-02	−1.5774e-02	1.1385e-01	1.0883e-02
5.4799e-03	−7.0102e-03	2.8499e-02	5.0742e-03

Table A4. *C* Matrix

1	0	0	0	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0	0	0
0	0	0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	0	0	1

Manuscript received Apr. 3, 1989, and revision received July 31, 1989.