Automatica, Vol. 26, No. 3, pp. 459–474, 1990
Printed in Great Britain.

0005-1098/90 $3.00 + 0.00
Pergamon Press plc

Survey Paper

Fault Diagnosis in Dynamic Systems Using Analytical and Knowledge-based Redundancy—A Survey and Some New Results*

PAUL M. FRANK†

A review of the different concepts of model-based fault detection and isolation in dynamic processes indicates that the practical application can be considerably improved by a systematic exploitation of the potentialities of robust residual generation.

Key Words—Failure detection; fault diagnosis; systems failure and recovery; reliability; sensor failures; actuator failures; redundancy; robustness; observers; expert systems.

Abstract—The paper reviews the state of the art of fault detection and isolation in automatic processes using analytical redundancy, and presents some new results. It outlines the principles and most important techniques of model-based residual generation using parameter identification and state estimation methods with emphasis upon the latest attempts to achieve robustness with respect to modelling errors. A solution to the fundamental problem of robust fault detection, providing the maximum achievable robustness by decoupling the effects of faults from each other and from the effects of modelling errors, is given. This approach not only completes the theory but is also of great importance for practical applications. For the case where the prerequisites for complete decoupling are not given, two approximate solutions—one in the time domain and one in the frequency domain—are presented, and the crossconnections to earlier approaches are evidenced. The resulting observer schemes for robust instrument fault detection, component fault detection, and actuator fault detection are briefly discussed. Finally, the basic scheme of fault diagnosis using a combination of analytical and knowledge-based redundancy is outlined.

1. *Introduction*

AUTOMATIC CONTROL systems are becoming more and more complex and the control algorithms more and more sophisticated. Consequently, there is a growing demand for fault tolerance which can be achieved not only by improving the individual reliabilities of the functional units but also by an efficient fault detection, isolation and accommodation (FDIA) concept.

* Received 20 February 1987; revised 29 August 1989; received in final form 14 September 1989. The original version of this paper was presented at the IFAC/IMAC/IFORS International Symposium on Advanced Information Processing in Automatic Control which was held in Nancy, France during July 1989. The Published Proceedings of this IFAC Meeting may be ordered from: Pergamon Press plc, Headington Hill Hall, Oxford OX3 0BW, U.K. This paper was recommended for publication in revised form by Editor K. J. Åström.

† Universität-GH-Duisburg, Fachgebiet Meß- und Regelungstechnik, F.R.G. This work was sponsored by the Deutsche Forschungsgemeinschaft and the Ministerium für Wissenschaft und Forschung des Landes Nordrhein-Westfalen.

This paper focuses on fault detection and isolation (FDI). In this context, a fault is understood as any kind of malfunction in the actual dynamic system, the plant, that leads to an unacceptable anomaly in the overall system performance. Such malfunctions may occur either in the sensors (instruments), or actuators, or in the components of the process. With respect to the different sectors where the faults can occur, one distinguishes between instrument fault detection (IFD), actuator fault detection (AFD), and component fault detection (CFD).

Over the last two decades, the basic research on FDI has gained increasing consideration world-wide. This development was (and still is) mainly stimulated by the above-mentioned trend of automation towards more complexity and the corresponding growing demand for higher availability and security of control systems. However, a strong impetus also comes from the side of modern control theory that has brought forth powerful techniques of mathematical modelling, state estimation and parameter identification that have been made feasible by the spectacular progresses of modern computer technology.

In the course of this development a novel philosophy for the FDI methodology has emerged and is increasingly discussed. It is based on the use of *analytical* (i.e. *functional*) rather than *physical* redundancy. This implies that the inherent redundancy contained in the static and dynamic relationships among the system inputs and measured outputs is exploited for FDI. In other words, one makes use of a *mathematical model* of the system or of parts of it.

Logically, there is some potential in using knowledge-based models instead of analytical models. This is the only way of FDI in all such cases where analytical models are not available. Therefore the knowledge-based approach may be looked upon as an alternative to the analytical model-based approach, or may complement it. The architecture of a combined analytical model-based and knowledge-based realtime diagnosis system will be outlined at the end of the paper.

Other than the physical redundancy approach that gets along with simple majority voting logic, the analytical redundancy approach and knowledge-based approach require advanced information processing techniques such as state estimation, parameter estimation, adaptive filtering, variable threshold logic, statistical decision theory, pattern recognition, heuristic reasoning, and various logical operations, all of which can be performed on a digital computer.

Various approaches to FDI using analytical redundancy have been reported in the last two decades. The number of different approaches, however, can be traced back to a few basic concepts. Among them are the detection filter (Beard, 1971; Jones, 1973), the innovation test using a single Kalman filter (Mehra and Peshon, 1971) or banks of Kalman filters or Luenberger observers (Clark *et al.*, 1975; Montgomery and Caglayan, 1974), the parity space approach (Deckert *et al.*, 1977), the parameter estimation technique (Kitamura, 1980) and expert system applications (see Tzafestas in the book by Patton *et al.*, 1989).

The appeal of the analytical redundancy approach lies in the fact that the existing redundancy can simply be evaluated by information processing under well-featured operating conditions (i.e. in the operation center) without need of additive physical instrumentation in the plant. However, there is a price to pay for this benefit which arises from the need of the mathematical model. Not only is there considerably more computational expenditure required for on-line modelling of the process; a much more serious problem is that of the sensitivity of the detection system with respect to modelling errors that are by no means avoidable in practice. Logically, the effect of modelling errors obscures the effect of faults and is therefore a source of false alarms. Hence, the sensitivity to modelling errors has become the key problem in the application of FDI methods based on analytical redundancy and deserves particular attention.

Even though the sensitivity problem was recognized early on (Deckert *et al.*, 1977; Clark, 1978a, b) effective methods for a systematic design of analytical-redundancy–FDI methods that are robust, i.e. insensitive or even invariant to modelling errors, have only been developed lately—see, e.g. Frank and Keller (1980), Watanabe and Himmelblau (1982), Frank and Keller (1984), Chow and Willsky (1984), Weiss *et al.* (1985), Emami-Naeini *et al.* (1986), Massoumnia (1986), Lou *et al.* (1986), Viswanadham and Srichander (1987), Patton *et al.* (1987), Wünnenberg and Frank (1987), Frank (1987a, b), Ge and Fang (1988) and Patton *et al.* (1989). The robustness issue will be emphasized in this paper.

For more details and earlier comprehensive surveys we refer to the articles of Willsky (1976), Walker (1983), Isermann (1984), Chow and Willsky (1984), Frank (1987a, b; 1988) and the books of Himmelblau (1978), Basseville and Benveniste (1985), Viswanadham *et al.* (1987), and Patton *et al.* (1989).

2. *Formulation of the fault detection and isolation problem*

In this section we outline the general procedure of FDI using analytical redundancy. Since the achievable quality of an FDI scheme mainly depends upon the quality of the model of the system, it is most important to start with a thorough and realistic specification of the given process. This will be the basis for the later fundamental solution of the FDI problem.

2.1. *System specification.* Consider a dynamic system with input vector, **u**, and output vector, **y**, possibly the plant in a feedback control system, as shown in Fig. 1. In general it consists of the actuators, the plant dynamics (components), and the sensors. For a realistic representation with respect to later use in the FDI task it is important to model all effects that can lead to alarms or false alarms. Such effects are:

(1) Faults in the *actuators,* or in the components of the plant dynamics, or in the sensors.
(2) *Modelling errors* between the actual system and its mathematical model.
(3) *System noise* and *measurement noise.*

This is illustrated in Fig. 1a. Figure 1b shows the simplified block representation where all the faults are composed to a fault vector **f** and all the other effects that obscure the fault detection to the so-called *vector of unknown inputs,* **d**.

Limiting our consideration to linear systems (for ease of explanation), the actual system may be given in continuous time by the state equations

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{E}\mathbf{d}(t) + \mathbf{K}\mathbf{f}(t) \quad (1)$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{F}\mathbf{d}(t) + \mathbf{G}\mathbf{f}(t) \quad (2)$$

where $\mathbf{x}$ is the $n \times 1$ state vector, $\mathbf{u}$ the $p \times 1$ known input vector, $\mathbf{y}$ the $q \times 1$ vector of measured outputs and **A**, **B**, **C** known matrices of appropriate dimensions. The term **Ed** models the unknown inputs to the actuators and to the dynamic process, **Kf** actuator and component faults, **Fd** the unknown

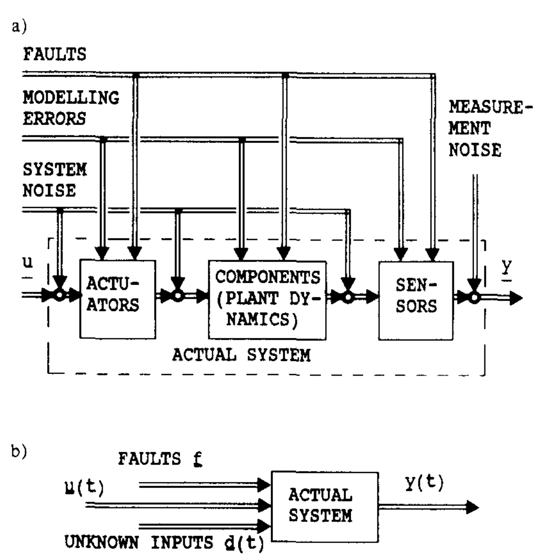

FIG. 1. System representation: (a) given situation, (b) simplified block representation.

inputs to the sensors, and **Gf** sensor faults. Notice that **A**, **B**, **C** are the nominal matrices of the system since the faults that are principally reflected in changes of **A**, **B**, **C**, as well as modelling errors, are considered by **f** and **d** associated with proper choices of **E**, **F**, **G**, **K**. Whilst these matrices are usually given, the modes (i.e. the evolutions) of **f** and **d** must generally be considered unknown.

The variety of fault modes that can occur may be classified as follows:

(1) Abrupt (sudden) faults, i.e. step-like changes.
(2) Incipient (slowly developing) faults, e.g. bias or drift.

Typically, *abrupt* faults play a role in safety-relevant systems where hard-failures have to be detected early enough so that catastrophic consequences can be avoided by early system reconfiguration. At the other end, *incipient* faults are of major relevance in connection with maintenance problems where early detection of worn equipment is required. In this case the faults are typically small and not as easy to detect, but the detection time is of minor importance and may therefore be large.

It should be noted that in all the cases where analytical models of the system are not available there is still some potential for fault diagnosis by the use of rule-based (qualitative) models that carry some kind of expert knowledge. For further discussion of the FDI approach based on *knowledge-based redundancy* we refer to Milne (1987) and Patton *et al.* (1989).

2.2. *The general procedure of model-based FDI.* The procedure of evaluation of the redundancy given by the mathematical model of the system, equations (1) and (2), can be roughly divided into the following two steps:

(1) *Generation* of so-called *residuals,* i.e. functions that are accentuated by the fault vector **f**.
(2) *Decision* and *isolation* of the faults (time, location, sometimes also type, size, and source).

The schematic conceptual structure of the FDI procedure using analytical redundancy (applied to a plant of a feedback control system) is illustrated in Fig. 2.

The analytical redundancy approach requires that the residual generator performs some kind of validation of the nominal relationships of the system,

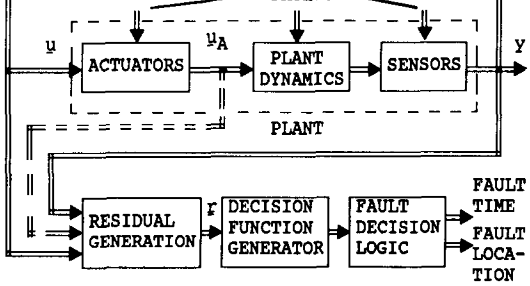

FIG. 2. Conceptual structure of FDI using analytical redundancy.

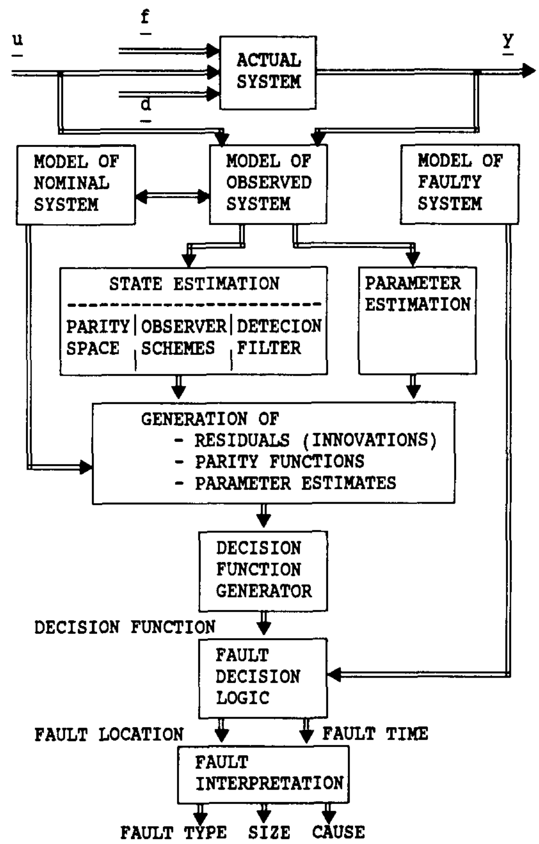

FIG. 3. General architecture of FDI based on analytical redundancy (Isermann, 1984).

using the actual input, **u**, and measured output, **y**. The redundancy relations to be evaluated can simply be interpreted as input–output relations of the dynamics of the system. Notice that the use of $\mathbf{u}_A$ instead of **u** may be an advantage for IFD and CFD when the actuators are highly nonlinear, because then the required system equations do not contain the actuator nonlinearities. If a fault occurs, the redundancy relations are no longer satisfied and a residual, $\mathbf{r} \neq \mathbf{0}$, occurs. The residual is then used to form appropriate decision functions. They are evaluated in the fault decision logic in order to monitor both the time of occurrence and location of the fault.

A more detailed structural diagram of the overall FDI procedure is depicted in Fig. 3. Note that for the residual generation three kinds of models are required: nominal, actual (observed) and that of the faulty system. In order to achieve a high performance of fault detection with low false alarm rate the nominal model should be tracked and updated by the observation model.

Basically, there are two different ways of generating fault-accentuated signals using analytical redundancy: be parity checks, observer schemes and detection filters on one side, all of them using state estimation techniques, and by parameter estimation on the other side. The resulting signals are used to form decision functions as, for example, norms or likelihood functions.

The basis for the decision on the occurrence of a fault is the fault *signature,* i.e. a signal that is obtained from some kind of faulty system model defining the effects associated with a fault. In most applications the FDI process is completed when the fault location and fault time are identified. In special cases, however, it may be desirable to get a deeper insight into the situation by knowing the fault type, size and cause which can be acquired by subsequent fault diagnosis. For this purpose deeper knowledge about the nature of the process such as, for example, the degree of ageing, the operational environment, used tools, history of operation and maintenance, fault statistics etc. is required. This task is therefore commonly tackled with the aid of an expert system.

3. *Residual generation*

3.1. *Problem statement.* Typically, the residual generation problem can be stated as follows. Consider a dynamic system with a known nominal mathematical model as, for example, represented by (1) and (2). Given the actual input vector, $\mathbf{u}(t)$, [or actuator output, $\mathbf{u}_A(t)$] and the measurement vector, $\mathbf{y}(t)$, suppose that a residual vector $\mathbf{r}(t)$ exists that carries information about a particular fault. Find an algorithm that generates $\mathbf{r}(t)$ when the fault has occurred, under the following conditions:
(1) The mode (time evolution) of the fault is unknown.
(2) The mathematical model of the nominal system is uncertain (with unknown tolerances).
(3) There is system noise and measurement noise (with unknown characteristics).
(4) The residual generation has to be performed within a specified time.

Of course, there are many modifications of the problem statement depending upon the given situation and the particular purpose of application.

3.2. *Conditions for the existence of a solution.* For the detectability and distinguishability of a fault the following conditions must hold:
(1) Knowledge of the "*normal*" behaviour, i.e. of the nominal model (**A**, **B**, **C**).
(2) Definitiveness of the "faulty" behaviour.
(3) Existence of analytical redundancy relations.
(4) Availability of at least one observation reflecting the fault.
(5) Satisfactory *reliability* of redundant information, i.e. invariance or at least robustness with respect to the unknown inputs.

4. *Methods of residual generation*

Looking at the literature one can see that there is a variety of different approaches to the problem of FDI using analytical redundancy. The great variety of proposed methods can, however, be brought down to a few basic concepts such as:
(1) the parity space approach;
(2) the dedicated observer approach and innovation-based approach;
(3) the fault detection filter approach;
(4) the parameter identification approach.

One should realize that there are some crossconnections between the first three approaches since they all end up with state estimators (observers or Kalman filters). In this section, the four approaches will be outlined.

4.1. *Parity space approach.* The early contributions to the parity space approach were made by Potter and Suman (1977), Desai and Ray (1981), and the group around Willsky (Chow and Willsky, 1984; Lou *et al.* 1986).

The key idea is to check the parity (consistency) of the mathematical equations of the system (analytical redundancy relations) by using the actual measurements. A fault is declared to have occurred once preassigned error bounds b_i are surpassed.

There are basically two forms of analytical redundancy relations:
(1) *direct redundancy*: relationships among instantaneous redundant sensor outputs (algebraic relations);
(2) *temporal redundancy*: dynamic relationships between sensor outputs and actuator inputs (differential or difference equations).

To outline the basic idea of the parity space methodology we consider first the simplified case of redundant measurements that may be obtained directly or from analytic sources (Desai and Ray, 1981). Let them be modeled by the algebraic measurement equation

$$\mathbf{y} = \mathbf{C}\mathbf{x} + \Delta\mathbf{y} \tag{3}$$

where $\mathbf{y}$ is the $(q \times 1)$ measurement vector, $\mathbf{C}$ the $(q \times n)$ measurement matrix of rank n, $\mathbf{x}$ the $(n \times 1)$ true measurement value, and $\Delta\mathbf{y}$ the $(q \times 1)$ error vector; $\Delta y_i > b_i$ defines a faulty operation indicated by the ith measured variable.

For detection of $\Delta\mathbf{y}$ the vector $\mathbf{y}$ can be combined to a set of linearly independent parity equations given by

$$\mathbf{p} = \mathbf{V}\mathbf{y} \tag{4}$$

where $\mathbf{p}$ is a $(q-n)$-dimensional parity vector. The $(q-n) \times q$ projection matrix $\mathbf{V}$ is now determined such that

$$\mathbf{V}\mathbf{C} = 0 \tag{5}$$

$$\mathbf{V}^{\mathrm{T}}\mathbf{V} = \mathbf{I}_q - \mathbf{C}(\mathbf{C}^{\mathrm{T}}\mathbf{C})^{-1}\mathbf{C}^{\mathrm{T}} \tag{6}$$

and

$$\mathbf{V}\mathbf{V}^{\mathrm{T}} = \mathbf{I}_{q-n}, \tag{7}$$

i.e. the rows of $\mathbf{V}$ are orthogonal, $\mathbf{V}$ is a null-space of $\mathbf{C}$. Hence

$$\mathbf{p} = \mathbf{V}\,\Delta\mathbf{y}. \tag{8}$$

This reveals that the parity equations contain only the errors due to the faults, independent of $\mathbf{x}$ which is not directly measured. Moreover, in the parity space, the columns of $\mathbf{V}$ define q distinct fault directions associated with each measurement. For example, the ith column of $\mathbf{V}$ determines the direction along which $\mathbf{p}$ lies if $\Delta\mathbf{y} = \Delta\mathbf{y}_i = [0 \cdots 0\ \Delta y_i\ 0 \cdots 0]^{\mathrm{T}}$. This ensures that a fault in measurement i implies a growth of $\mathbf{p}$ in the ith direction.

Obviously, the q-dimensional residual vector $\mathbf{r} = \mathbf{y} - \mathbf{C}\hat{\mathbf{x}}$, where $\hat{\mathbf{x}} = (\mathbf{C}^T\mathbf{C})^{-1}\mathbf{C}^T\mathbf{y}$ is the least squares estimate of $\mathbf{x}$, is related to the parity vector $\mathbf{p}$ as

$$\mathbf{r} = \mathbf{V}^T\mathbf{p} \tag{9}$$

In these terms the FDI problem can be formulated as follows. Given q redundant measurements $y_1, \ldots, y_q$ of a process variable, and symmetric error bounds $b_1, \ldots, b_q$ characterizing the faulty behavior:

(1) Find an estimate $\hat{\mathbf{x}}$ of the process variable from a most consistent subset of measurements.

(2) Identify the faulty measurement by parity checks.

Obviously, to detect a single fault among p components at least $p-1$ parity relations are required.

This concept was generalized by Chow and Willsky (1984) and Lou *et al.* (1986) for the case of using the temporal redundancy relations of the dynamic system. Suppose the system is given by the linear discrete state equations

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) \tag{10}$$

$$y(k) = \mathbf{C}\mathbf{x}(k) \tag{11}$$

where $\mathbf{x}$ is the $n \times 1$ state vector, $\mathbf{u}$ the $p \times 1$ actuator input vector, $\mathbf{y}$ the $q \times 1$ sensor output vector, and $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ matrices of appropriate dimensions. The redundancy relations are now specified mathematically as follows. Let us first define the subspace of $(s+1)q$ dimensional vectors $\mathbf{v}$ by

$$P = \left\{ \mathbf{v} \mid \mathbf{v}^T \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^s \end{bmatrix} = 0 \right\} \tag{12}$$

which is called the parity space of order s. Every vector $\mathbf{v}$ can be used at any time k for a parity check:

$$r(k) = \mathbf{v}^T \left[\begin{bmatrix} \mathbf{y}(k-s) \\ \vdots \\ \mathbf{y}(k) \end{bmatrix} - \boldsymbol{H} \begin{bmatrix} \boldsymbol{u}(k-s) \\ \vdots \\ \mathbf{u}(k) \end{bmatrix} \right] \tag{13}$$

with

$$\mathbf{H} = \begin{bmatrix} 0 & & & & & \\ \mathbf{CB} & 0 & & & \mathbf{0} & \\ \mathbf{CAB} & \mathbf{CB} & 0 & & & \\ \vdots & & & & & \\ \mathbf{CA}^{s-1}\mathbf{B} & \cdots & \cdots & \mathbf{CAB} & \mathbf{CB} & 0 \end{bmatrix}. \tag{14}$$

Substituting the state equations (10), (11) in $r(k)$ yields

$$r(k) = \mathbf{v}^T \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^s \end{bmatrix} \mathbf{x}(k-s). \tag{15}$$

Due to the above definition of $\mathbf{v}$, $r(k)$ is zero if no faults occur. From this it becomes evident that a redundancy relation is simply an input–output model for a part of the dynamics of the system. In other words, instead of checking the consistency of the overall mathematical model one restrict oneself to checking individual relations that are part of the model. This allows one to select the most reliable relations and thus create robustness of the FDI procedure.

For a geometrical interpretation of the above defined parity checks consider the model (10), (11) and let Z denote the range of the matrix in (12). Then the parity space P is the orthogonal complement of Z, and a complete set of parity checks of order s is given by the orthogonal projection of the vector of input-adjusted observations onto P [given by the quantity in brackets $[\cdot]$ in equation (13)].

Note that the parity space approach leads to a special type of observer for fault detection, namely the so-called *dead-beat observer*. This holds not only for SISO systems as was shown by Massoumnia (1986) but also for MIMO systems (Frank *et al.* in the book by Patton *et al.* 1989).

4.2. *Dedicated observer approach.* From the above it becomes apparent that the closed-loop parity space approach leads to the concept of state estimation. Independent from this, many authors have approached the FDI problem by directly starting with single or banks of Luenberger observers or Kalman filters; see, for example, Clark *et al.* (1975), Willsky (1976), Clark (1978a, b), Frank and Keller (1980), Frank (1978b; 1988), Mehra and Peshon (1971) and Willsky (1976).

The basic idea of the observer approach is to reconstruct the outputs of the system from the measurements or subsets of the measurements with the aid of observers or Kalman filters using the estimation error or innovation, respectively, as a residual for the detection and isolation of the faults. It is well known from observer theory that for state estimation one can use linear or nonlinear, full or reduced-order state observers in the deterministic case or Kalman filters in the stochastic case when noise has to be considered.

The fundamental configuration of a linear full order state estimator is shown in Fig. 4. Notice that the full order observer simply consists of a parallel model of the process with a feedback of the estimation error, $\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}}$. Though, in principle, the open-loop model would also do, the feedback is important for several

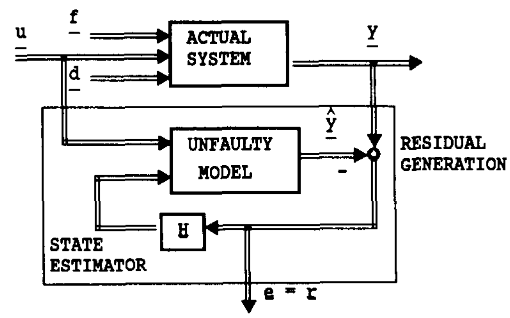

FIG. 4. Basic configuration of residual generation through state estimation.

reasons:

(1) to compensate for differences in the initial conditions;
(2) to stabilize the parallel model in case of an unstable system;
(3) to provide freedom for the *design* of the filter, for example, to decouple the effects of faults from the effects of other faults or unknown inputs.

In the case of a linear process with the state equations (1), (2), the state, $\hat{\mathbf{x}}$, and output, $\hat{\mathbf{y}}$, of a full-order observer are governed by the equations

$$\dot{\hat{\mathbf{x}}} = (\mathbf{A} - \mathbf{HC})\hat{\mathbf{x}} + \mathbf{Bu} + \mathbf{Hy} \tag{16}$$

$$\hat{\mathbf{y}} = \mathbf{C}\hat{\mathbf{x}} \tag{17}$$

where $\mathbf{H}$ denotes the feedback gain matrix that has to be chosen properly to achieve a desired performance of the observer. With (1), (2), (16), (17) the relations for the state estimation error, $\underline{\epsilon} = \mathbf{x} - \hat{\mathbf{x}}$, and the output estimation error, $\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}}$, become

$$\dot{\underline{\epsilon}} = (\mathbf{A} - \mathbf{HC})\underline{\epsilon} + \mathbf{Ed} + \mathbf{Kf} - \mathbf{HFd} - \mathbf{HGf} \tag{18}$$

$$\mathbf{e} = \mathbf{C}\underline{\epsilon} + \mathbf{Fd} + \mathbf{Gf}. \tag{19}$$

It is seen from equations (18), (19) that the output estimation error, $\mathbf{e}$, is a function of $\mathbf{f}$ and $\mathbf{d}$ (but not of $\mathbf{u}$). Hence, $\mathbf{e}$ can be used as the residual, $\mathbf{r}$, for the purpose of detection and isolation of the fault. When no fault occurs, i.e. $\mathbf{f} = \mathbf{0}$, then $\mathbf{r}$ only will be influenced by the unknown input, $\mathbf{d}$. When, however, $\mathbf{f} \neq \mathbf{0}$, $\mathbf{r}$ will be increased. Thus, a fault can be detected by checking the increment of $\mathbf{r}$ caused by $\mathbf{f}$. In the simplest case this can be done by a threshold logic (Clark *et al.*, 1975; Clark, 1978a, b). To avoid false alarms the threshold must be chosen larger than zero, though this reduces the sensitivity to faults. Ways to overcome this difficulty will be discussed later. In a similar way one can generate residuals using reduced-order or nonlinear estimators (Clark, 1978a, b; Frank, 1987a, b).

The basic task of the design of state estimators for FDI is to optimize them by a proper choice of the feedback gain matrix $\mathbf{H}$. Before addressing this problem we give a brief outline of the different configurations of observer-based FDI schemes.

The most simple configuration used for instrument fault detection is a *single estimator* (*observer or Kalman filter*), where a single full or reduced-order estimator is driven by *only one* (the most reliable) sensor output, and the full output is reconstructed (Clark, 1978a). The comparison of the actual output, $\mathbf{y}$ with the estimated output, $\hat{\mathbf{y}}$, using a threshold logic allows, in principle, a unique detection and isolation of a single faulty instrument.

Mehra and Peshon (1971) use a *single Kalman* filter driven by the *full output* vector as shown in Fig. 5 and

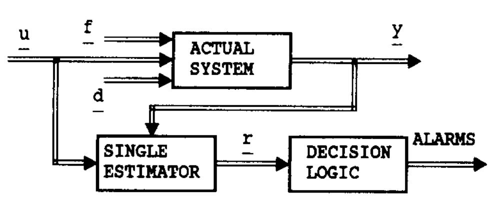

FIG. 5. IFD using a single estimator.

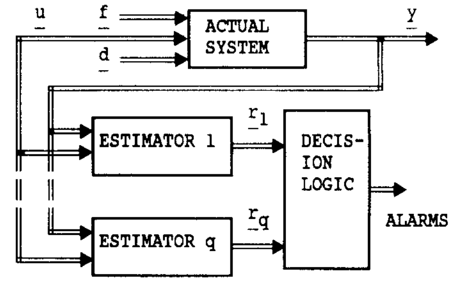

FIG. 6. IFD using a bank of estimators (estimator scheme).

make use of the fact that the innovation is white noise with zero mean and known covariance when no fault occurs. The occurrence of a fault is monitored by *statistical innovation tests* of whiteness, mean and covariance. The fault isolation is carried out on the basis of different fault hypotheses. The *multiple hypotheses testing* can, for example, be carried out using Bayesian decision theory (Willsky, 1976).

More flexibility in the isolation of actuator, component or sensor faults can be achieved by using an estimator *scheme*, i.e. a *bank* of estimators driven by the actual output vector, $\mathbf{y}$, or subsets of $\mathbf{y}$ (Fig. 6.) A popular approach is based on *multiple hypotheses testing* (Willsky, 1976). In this case each of the estimators is designed for a different fault hypothesis (for example: H_0: no fault; H_1: bias in sensor 1; H_2: bias in sensor 2; H_3; zero output in sensor 1, etc.). The hypotheses are then tested in terms of likelihood functions using, for example, Bayesian decision theory (Wilbers and Speyer, 1989).

A well known approach for instrument fault detection is to assign a dedicated estimator to each of the sensors. In the *dedicated observer scheme* proposed for IFD by Clark (1978b) each estimator is driven by a different single sensor output and the complete output vector $\mathbf{y}$ or, if this is not possible, as many components of $\mathbf{y}$ as possible, are estimated. With this scheme multiple simultaneous faults can be principally detected and isolated by checking properly structured sets of estimation errors, for example, with the aid of a threshold logic. If, for example, a certain sensor fails, then the related output estimate reconstructed by the corresponding estimator will be erroneous which can be identified by the logic.

An alternative version, the so-called *generalized observer scheme* (Frank, 1987a, b) provides that an estimator dedicated to a certain sensor is driven by all outputs except that of the respective sensor. This IFD scheme allows one to detect and isolate only a single fault in any of the sensors, however, with increased robustness with respect to unknown inputs. This problem will be addressed more systematically in Section 6.

Instead of using pure software redundancy one can also combine software with hardware redundancy with the benefit of saving hardware expenditure. In the most common practice of a *duplex sensor system* (Fig. 7), one uses two identical sets of instruments, each set being supervised by an IFD scheme of one of the above-mentioned types (Deckert *et al.*, 1977; Onken and Stuckenberg, 1979). Once a fault in one of the

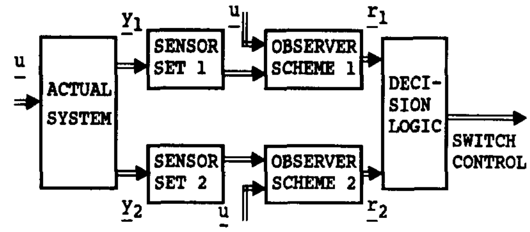

FIG. 7. Duplex sensor system with analytical redundancy.

sensors has occurred, it is detected with the aid of the observer schemes and the system is then switched to the healthy sensor. The motivation for using two like sensors is to detect the occurrence of a fault by hardware redundancy and only perform the isolation task by analytical redundancy. This saves computational burden, increases the reliability and allows the analytical redundancy tests to be triggered. This concept has already gained some practical importance for IFD in aircrafts (Labarrère, 1987).

Little work has been done so far on the study of FDI schemes using *nonlinear estimators* though the nonlinear estimator theory is well settled. Logically, when the system under consideration is nonlinear, the model in the observer should also be nonlinear in order to avoid modelling errors arising from linearization. This leads to the concept of FDI using nonlinear state estimators which will now be briefly discussed.

Consider the nonlinear system given by

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}); \ \mathbf{x}(0) = \mathbf{x}^0 \tag{20}$$

$$\mathbf{y} = \mathbf{c}(\mathbf{x}, \mathbf{u}) \tag{21}$$

with $\mathbf{x}$, $\mathbf{u}$, $\mathbf{y}$ having the same dimensions as in (1), (2). The nonlinear state estimator equation is then, by definition,

$$\dot{\hat{\mathbf{x}}} = \hat{\mathbf{f}}(\hat{\mathbf{x}}, \mathbf{u}, \mathbf{y}), \quad \hat{\mathbf{x}}(0) = \mathbf{x}^0 \tag{22}$$

and the state estimation error $\underline{\epsilon} = \mathbf{x} - \hat{\mathbf{x}}$ becomes

$$\dot{\underline{\epsilon}} = \mathbf{f}(\mathbf{x}, \mathbf{u}) - \hat{\mathbf{f}}(\hat{\mathbf{x}}, \mathbf{u}, \mathbf{y}). \tag{23}$$

Under certain conditions one can approximate equation (22) such that it becomes

$$\dot{\hat{\mathbf{x}}} = \mathbf{f}(\hat{\mathbf{x}}, \mathbf{u}) + \mathbf{H}(\hat{\mathbf{x}}, \mathbf{u})(\mathbf{y} - \hat{\mathbf{y}}); \quad \hat{\mathbf{x}}(0) = \hat{\mathbf{x}}^0 \tag{24}$$

$$\hat{\mathbf{y}} = \mathbf{c}(\hat{\mathbf{x}}, \mathbf{u}) \tag{25}$$

where

$$\mathbf{H}(\hat{\mathbf{x}}, \mathbf{u}) = \left.\frac{\partial \hat{\mathbf{f}}}{\partial \mathbf{y}}\right|_{\hat{\mathbf{x}}, \mathbf{u}} \tag{26}$$

is a time-variant observer gain matrix.

If system noise $\underline{\nu}(t)$ and modelling errors $\Delta\mathbf{f}(t)$ are present, the state estimation error equation becomes

$$\dot{\underline{\epsilon}} = \left[\frac{\partial \mathbf{f}}{\partial \mathbf{x}} - \mathbf{H}(\mathbf{x}, \mathbf{u})\frac{\partial \mathbf{c}}{\partial \mathbf{x}}\right]_{\hat{\mathbf{x}}, \mathbf{u}} \underline{\epsilon} + \Delta\mathbf{f} + \underline{\nu} \tag{27}$$

from which the output estimation error $\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}}$ can be calculated. Considering measurement noise, $\underline{\mu}(t)$, and sensor faults $\Delta\mathbf{k}(t)$, one obtains

$$\mathbf{e} = \mathbf{c}(\mathbf{x}, \mathbf{u}) - \mathbf{c}(\hat{\mathbf{x}}, \mathbf{u}) + \Delta\mathbf{k} + \underline{\mu}. \tag{28}$$

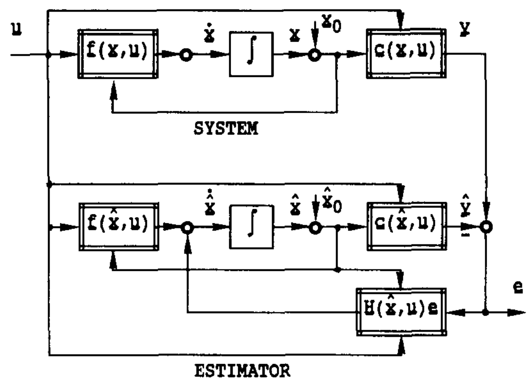

FIG. 8. Residual generation for a nonlinear system using a nonlinear observer.

The structural diagram of the resulting nonlinear estimator is illustrated in Fig. 8.

Note that in many practical situations, stability of the observer can be maintained using a constant feedback gain matrix rather than $\mathbf{H}(\hat{\mathbf{x}}, \mathbf{u})$. This simplifies the implementation of the FDI scheme; however, it obscures the understanding of its behaviour and complicates assessment of its overall stability. For more details and further references see Frank (1987a).

An elegant approach to the FDI of a certain class of nonlinear systems is that of using nonlinear decoupling (Wünnenberg, 1990; Wünnenberg and Frank, 1990). This leads to a state estimation error equation that is linear in the estimation error space and therefore easy to evaluate. The prerequisites defining the class of nonlinear systems to which this method is applicable is that the system can be described by the following type of state equations

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}(\mathbf{y}, \mathbf{u}) + \mathbf{E}\mathbf{d} + \mathbf{K}(\mathbf{x})\mathbf{f} \tag{29}$$

$$\mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{G}(\mathbf{x})\mathbf{f} \tag{30}$$

with the appropriate matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{E}$, $\mathbf{K}$, $\mathbf{C}$, $\mathbf{G}$. Notice that characteristic of this type of system is that the system nonlinearities of the state equations can be expressed as an explicit function of the input, $\mathbf{u}$, and measured output, $\mathbf{y}$ and that the measurement is only corrupted by the faults and not by unknown inputs.

The corresponding fault detection observer is then given by

$$\dot{\mathbf{z}} = \mathbf{R}\mathbf{z} + \mathbf{J}(\mathbf{y}, \mathbf{u}) + \mathbf{S}\mathbf{y} \tag{31}$$

$$\mathbf{r} = \mathbf{L}_1\mathbf{z} + \mathbf{L}_2\mathbf{y}. \tag{32}$$

Assuming that, in the unfaulty case, $\mathbf{z}$ results from the linear transformation of $\mathbf{x}$,

$$\mathbf{z} = \mathbf{T}\mathbf{x}, \tag{33}$$

the estimation error equation becomes

$$\dot{\mathbf{e}} = \dot{\mathbf{z}} - \mathbf{T}\dot{\mathbf{x}} = \mathbf{R}\mathbf{z} + \mathbf{J}(\mathbf{y}, \mathbf{u}) + \mathbf{S}\mathbf{y} - \mathbf{T}\mathbf{A}\mathbf{x} - \mathbf{T}\mathbf{B}(\mathbf{y}, \mathbf{u}) - \mathbf{T}\mathbf{E}\mathbf{d} - \mathbf{T}\mathbf{K}(\mathbf{x})\mathbf{f}. \tag{34}$$

This leads to the equations

$$\begin{aligned} \mathbf{T}\mathbf{A} - \mathbf{R}\mathbf{T} &= \mathbf{S}\mathbf{C} \\ \mathbf{T}\mathbf{E} &= \mathbf{0} \\ \mathbf{J}(\mathbf{y}, \mathbf{u}) &= \mathbf{T}\mathbf{B}(\mathbf{y}, \mathbf{u}) \\ \mathbf{L}_1\mathbf{T} + \mathbf{L}_2\mathbf{C} &= \mathbf{0} \end{aligned} \tag{35}$$

from which the matrices of $\mathbf{T}$, $\mathbf{R}$, $\mathbf{J}$, $\mathbf{S}$, $\mathbf{L}_1$, $\mathbf{L}_2$ can be determined.

Once $\mathbf{T}$ is found, the rest of the problem is easy to solve. The solution (for $\mathbf{T}$) can be obtained, for example, using the unknown input observer approach (Patton *et al.*, 1989) to be discussed in Section 5. This also applies to the determination of the approximate solution for the case that the above conditions cannot be fulfilled exactly (Wünnenberg, 1990).

4.3. *Fault detection filter* (*FDF*). The fault detection filter (or fault sensitive filter) is a full-order state estimator with a special choice of $\mathbf{H}$. It was first proposed by Beard (1971) and Jones (1973).

To describe the underlying idea following Wilbers and Speyer (1989), let us start with the system state equations in the form

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{k}_i f_i(t) \tag{36}$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \bar{\mathbf{k}}_j \bar{f}_j(t) \tag{37}$$

where $\mathbf{k}_i$ is an $n \times 1$ design fault direction, $i = 1, 2, \ldots, r$, r is the number of fault directions, and $f_i(t)$ is an arbitrary scalar function of time. When no fault occurs, $f_i = 0$. The fault directions $\mathbf{k}_i$ can be used to model actuator and component faults. For example, the ith actuator fault can be represented by $\mathbf{k}_i = \mathbf{b}_i$ where $\mathbf{b}_i$ is the ith column in $\mathbf{B}$. Accordingly, one can model the directions and modes of sensor faults by $\bar{\mathbf{k}}_j$ and $\bar{f}_j(t)$, respectively, where $\bar{\mathbf{k}}_j$ is a $q \times 1$ unit vector associated with a fault in the jth sensor.

The corresponding observer equations become

$$\dot{\hat{\mathbf{x}}} = (\mathbf{A} - \mathbf{H}\mathbf{C})\hat{\mathbf{x}} + \mathbf{B}\mathbf{u} + \mathbf{H}\mathbf{y} \tag{38}$$

$$\hat{\mathbf{y}} = \mathbf{C}\hat{\mathbf{x}} \tag{39}$$

where $\mathbf{H}$ denotes the feedback gain matrix. The $\mathbf{H}$ matrix is now chosen such that the residual vector (i.e. the output error) $\mathbf{r} = \mathbf{y} - \hat{\mathbf{y}}$ has certain directional properties at the occurrence of a certain fault. If an *actuator* or *component fault* occurs then the equation for the state error vector $\underline{\epsilon} = \mathbf{x} - \hat{\mathbf{x}}$ becomes

$$\dot{\underline{\epsilon}} = (\mathbf{A} - \mathbf{H}\mathbf{C})\underline{\epsilon} + \mathbf{k}_i f_i \tag{40}$$

$$\mathbf{r} = \mathbf{C}\underline{\epsilon}. \tag{41}$$

For a *sensor fault* we have

$$\dot{\underline{\epsilon}}_j = (\mathbf{A} - \mathbf{H}\mathbf{C})\underline{\epsilon}_j + \mathbf{h}_j \bar{f}_j \tag{42}$$

$$\mathbf{r}_j = \mathbf{C}\underline{\epsilon}_j + \bar{\mathbf{k}}_j \bar{f}_j \tag{43}$$

where $\mathbf{h}_j$ is the jth column of the detection filter gain matrix. This system can be replaced by

$$\dot{\underline{\epsilon}}_j = (\mathbf{A} - \mathbf{H}\mathbf{C})\underline{\epsilon}_j - \mathbf{k}_j^* \bar{f}_j + \mathbf{f}_j^* \tag{44}$$

$$\mathbf{r}_j = \mathbf{C}\underline{\epsilon}_j \tag{45}$$

where

$$\mathbf{k}_j^* = \mathbf{A}\mathbf{f}_i^* - \alpha\mathbf{f}_j^* \tag{46}$$

with an arbitrary scalar α. $\mathbf{f}_j^*$ is the fault direction associated with the ith sensor such that $\mathbf{C}\mathbf{f}_j^* = \bar{\mathbf{k}}_j$.

As can be seen, the residual of the detection filter can be made unidirectional in the case of actuator and component faults, but can only be made to lie in a plane for a sensor fault. $\mathbf{H}$ is chosen such that the actuator or component fault residuals come to lie in the direction of $\mathbf{C}\mathbf{k}_i$, and sensor fault residuals in the plane in which $\mathbf{C}\mathbf{f}_j^*$ and $\mathbf{C}\mathbf{k}_j^*$ lie.

In conclusion, typical for the fault detection filter is the proper choice of $\mathbf{H}$ so that the residual $\mathbf{r}$ due to a *particular fault*, f_i (component of $\mathbf{f}$) is constrained to a single direction or plane in the residual space independent of the mode of f_i. This is often not possible unless the state $\mathbf{x}$ is accordingly enlarged. Since the important information for the fault detection is in the *direction* of the residual rather than in its time function, the use of a fault detection filter does not require the knowledge of the fault mode. Hence, a fault is detected when one or more of the residual projections along the known fault direction or in the known fault plane are sufficiently large.

It is a most appealing feature of the fault detection filter that the residual *direction* for a fault is not affected by the fault *mode* $f_i(t)$ (i.e. the size or time history). Notice, however, that unknown inputs (including parameter variations) have not been considered in the context of detection filters up to now. Hence, this approach does not account for the effects of disturbances, parameter variations or measurement noise and therefore requires precise modelling (Wilbers and Speyer, 1989).

To complete our consideration of detection filters let us point to a simple method that has long been used in connection with state augmentation. The idea is to interpret a parameter, that reflects a possible fault, as a state variable and include it in the state vector, thus increasing the order of the system. If then the augmented state is observed by an (augmented) observer, the fault reflected in the corresponding parameter can be detected and isolated. This leads to a nonlinear extended Kalman filter which can be seen as a combination of state estimation and (least square) parameter identification. This method of state augmentation is well known in the literature but has not been widely used for fault detection yet.

4.4. *Parameter identification approach*. This is an alternative approach to the above described methods based on state estimation. It makes use of the fact that faults of a dynamic system are reflected in the physical parameters as, for example, friction, mass, viscosity, resistance, capacitance, inductance, etc. The idea of the parameter identification approach is to detect the faults via estimation of the parameters of the mathematical model due to the following procedure (Isermann, 1984):

(1) Choice of a parametric model of the system. Normally one uses linear models with lumped parameters with input/output differential equations of the form

$$a_n y^{(n)}(t) + \cdots + a_1 \dot{y}(t) + y = b_0 u(t) + \cdots + b_m u^{(m)}(t). \tag{47}$$

(2) Determination of the relationships between the model parameters θ_i and the physical parameters p_i

$$\underline{\theta} = \mathbf{f}(\mathbf{p}). \tag{48}$$

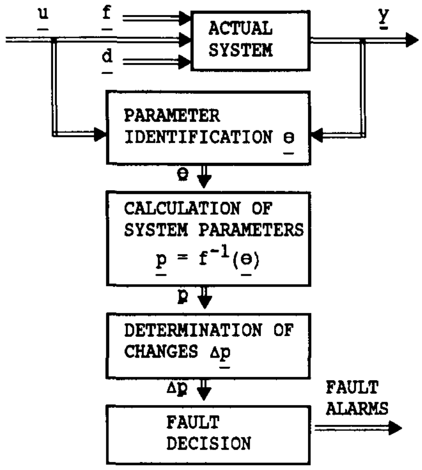

FIG. 9. Parameter identification approach to fault detection.

(3) Identification of the model parameter vector $\underline{\theta}$ using the input **u** and output **y** of the actual system (or of a system component).
(4) Determination of the physical parameter vector

$$\mathbf{p} = \mathbf{f}^{-1}(\underline{\theta}). \tag{49}$$

(5) Calculation of the vector of deviations, $\Delta\mathbf{p}$, from its nominal value taken from the nominal model.
(6) Decision on a fault by exploiting the relationships between faults and changes in the physical parameters, Δp_i.

The procedure is illustrated in Fig. 9. This approach may be particularly useful for the detection of incipient faults. A detailed description is given by Isermann (1984), and see also Patton *et al.* (1989).

5. *Increasing the robustness to unknown inputs*

According to the task definition of an FDI system given earlier the following two requirements are of primary importance:
(1) Distinguishability between different faults.
(2) Robustness with respect to unknown inputs (including the modelling errors).

In this section we focus our attention on the robustness problem.

5.1. *State of the art.* The FDI concepts using analytical redundancy principally suffer from a fundamental practical limitation in that the system model on which the redundancy is based is never known exactly. The consequence is that the actual system outputs, y_i, will not match the model outputs, $\hat{y}_i$, even when there are no faults present in the system. Thus the residual, $\mathbf{r} = \mathbf{y} - \hat{\mathbf{y}}$, will be nonzero in general and one is forced to use thresholds to distinguish a fault. The crux with thresholds is that they not only reduce the sensitivity to faults, but also vary with the input signal of the actual system and the magnitude and nature of the system disturbances. Choosing the threshold too low increases the rate of false alarms; choosing it too large reduces the net effect of fault detection. There is therefore a strong motivation to reduce the sensitivity of the residual with respect to modelling errors.

The robustness problem has been recognized early on and several approaches to increase the robustness of FDI schemes have been suggested over the years. Frank and co-workers have developed state estimator design techniques for robust residual generation, the results of which are summarized in Frank (1978a, b) and in the book by Patton *et al.* (1989). Other significant contributions to the robust observer design for FDI were made by Watanabe and Himmelblau (1982), Massoumnia (1986), Patton *et al.* (1987, 1989), Ge and Fang (1988) and others. In Chow and Willsky (1984), Lou *et al.* (1986), and Weiss *et al.* (1985) the robustness problem was primarily tackled from the parity space point of view. From this perspective the residual of an estimator can be viewed as the most general parity function containing the complete set of redundancy relations. The underlying idea of robustness generation is to utilize only those redundancy relations that are most reliable. Procedures for finding optimal solutions are given in Chow and Willsky (1984) and Lou *et al.* (1986).

On the other hand, there are currently several approaches to increase the FDI robustness either by a proper choice of the threshold (Emami-Naeini *et al.*, 1986) or by making the threshold adaptive to the input, as proposed by Clark in the book by Patton *et al.* (1989).

We shall now look more closely at the generation of robustness in the
(1) parameter estimation approach;
(2) state estimation approach.

5.2. *Robust parameter estimation approach.* Robustness to unmodelled dynamics with the aid of parameter estimation can be achieved by the following method. Suppose the parameter vector $\underline{\theta}$ of the actual dynamic system is determined by a standard least squares estimation using a model with undermodelled dynamics. $\underline{\theta}$ may take the value $\underline{\theta}_n$ in the nominal case, and $\underline{\theta}_f$ after a fault has occurred. The estimates of $\underline{\theta}_n$ and $\underline{\theta}_f$ are denoted as $\hat{\underline{\theta}}_n$ and $\hat{\underline{\theta}}_f$, respectively.

The fault detection problem consists of the decision between the following two hypotheses:

$$\begin{aligned} &\mathrm{H}_0\colon\ \underline{\theta}_n = \underline{\theta}_f \text{ (no fault)} \\ &\mathrm{H}_1\colon\ \underline{\theta}_n \neq \underline{\theta}_f \text{ (fault)}. \end{aligned} \tag{50}$$

In a first step, the expected difference between $\underline{\theta}_n$ and $\underline{\theta}_f$ under H_0 is tested in terms of the covariance, i.e.

$$\mathbf{S} = \mathrm{Cov}\,(\hat{\underline{\theta}}_n - \hat{\underline{\theta}}_f) \quad \text{for} \quad \underline{\theta}_n = \underline{\theta}_f. \tag{51}$$

The robust fault detection is now based on the idea of comparing the observed value of $(\hat{\underline{\theta}}_n - \hat{\underline{\theta}}_f)(\hat{\underline{\theta}}_n - \hat{\underline{\theta}}_f)^{\mathrm{T}}$ with its expected value, i.e. the covariance **S**.

The test can be performed using different measures. One possible measure is

$$r = (\hat{\underline{\theta}}_n - \hat{\underline{\theta}}_f)^{\mathrm{T}}\mathbf{S}^{-1}(\hat{\underline{\theta}}_n - \hat{\underline{\theta}}_f). \tag{52}$$

This is the solution of a minimization problem where the "robust" parameters have the dominant influence.

As an alternative expression for the residual generation one may use

$$r = \frac{|\hat{\theta}_n^i - \hat{\theta}_f^i|}{C_{ii}} \tag{53}$$

which is executed for all i elements $\hat{\theta}_n^i$ and $\hat{\theta}_f^i$ of $\hat{\theta}_n$ and $\hat{\theta}_f$, respectively. Finally, in either case the residual is compared with a fixed threshold, λ, and the H_0 hypothesis is accepted if $r \leq \lambda$.

Notice, however, that this methodology does not include maximizing the sensitivity with respect to the parameter variations to be detected. Such a method has not yet been proposed.

5.3. *Robust observer schemes.* Clearly, the key problem of designing *robust observer* schemes for FDI must primarily concern the stage of residual generation. The most demanding approach to the design of robust observers consists of choosing the feedback matrix $\mathbf{H}$ of the observer such that
(1) the fault effects are decoupled from each other;
(2) the residual becomes invariant to the vector of unknown inputs $\mathbf{d}$.

This is a generalization of the traditional fault detection filter approach in the sense that the fault effects on the residuals are not only decoupled from each other but also from unknown inputs. Several authors look upon this problem from the viewpoint of eigenstructure assignment (Patton *et al.*, 1989; Wilbers and Speyer, 1989). A systematic solution to this problem using the *unknown input observer* (UIO) approach was given by Wünnenberg and Frank (1987), Viswanadham and Srichander (1987) and Ge and Fang (1988).

In the following we will outline the exact solution of this problem as well as two approximations: one in the time domain and one in the frequency domain. Moreover, it will be shown that this approach is a unified one that contains the observer-based methods described above as special cases.

5.4. *A unified approach to robust residual generation using unknown input observers.* Let us assume that the mathematical model (1), (2) of the system be given in discrete form

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{E}\mathbf{d}_k + \mathbf{K}\mathbf{f}_k; \quad \mathbf{x}_{k=0} = \mathbf{x}_0 \tag{54}$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{F}\mathbf{d}_k + \mathbf{G}\mathbf{f}_k. \tag{55}$$

Recall that this representation includes all kinds of faults, namely component faults and actuator faults as modelled by the term $\mathbf{K}\mathbf{f}_k$, as well as sensor faults modelled by $\mathbf{G}\mathbf{f}_k$. This implies that the approach is applicable to component fault detection (CFD), actuator fault detection (AFD), and instrument fault detection (IFD).

The corresponding unknown input observer (UIO) that meets the robustness requirements as defined above must obey the following equation:

$$\mathbf{z}_{k+1} = \mathbf{R}\mathbf{z}_k + \mathbf{S}\mathbf{y}_k + \mathbf{J}\mathbf{u}_k \quad \mathbf{z}_{k=0} = \mathbf{z}_0 \tag{56}$$

with the residual

$$\mathbf{r}_k = \mathbf{L}_1\mathbf{z}_k + \mathbf{L}_2\mathbf{y}_k \tag{57}$$

and $\mathbf{r}_k$ having the following properties

$$\lim_{k\to\infty} \mathbf{r} = 0 \text{ for all } \mathbf{u} \text{ and } \mathbf{d} \text{ and for all initial conditions } \mathbf{x}_0 \text{ and } \mathbf{z}_0 \text{ and } \mathbf{f}_k = \mathbf{0} \tag{58}$$

$$\mathbf{r}_k \neq \mathbf{0} \quad \text{if } \mathbf{f}_k \neq \mathbf{0}. \tag{59}$$

In addition, the states $\mathbf{z}_k$ of the UIO are linear combinations of the system states according to

$$\mathbf{z}_k = \mathbf{T}\mathbf{x}_k \tag{60}$$

in the unfaulty case after the response to unlike conditions has died out. A necessary and sufficient condition for the existence of such an "ideal" UIO is given by Frank *et al.* in the book by Patton *et al.* (1989).

The corresponding equation for the estimation error of the above observer becomes

$$\mathbf{e}_{k+1} = \mathbf{z}_{k+1} - \mathbf{T}\mathbf{x}_{k+1} = \mathbf{R}\mathbf{z}_k + \mathbf{S}\mathbf{y}_k + \mathbf{J}\mathbf{u}_k - \mathbf{T}\mathbf{A}\mathbf{x}_k - \mathbf{T}\mathbf{B}\mathbf{u}_k - \mathbf{T}\mathbf{E}\mathbf{d}_k - \mathbf{T}\mathbf{K}\mathbf{f}_k. \tag{61}$$

To meet the robustness requirements, the following set of equations has to be solved:

$$\mathbf{TA} - \mathbf{RT} = \mathbf{SC} \tag{62}$$

$$\mathbf{J} = \mathbf{TB} \tag{63}$$

$$\mathbf{TE} = \mathbf{0} \tag{64}$$

$$\mathbf{SF} = \mathbf{0} \tag{65}$$

$$\mathbf{SG} \neq \mathbf{0} \tag{66}$$

$$\mathbf{TK} \neq \mathbf{0}. \tag{67}$$

Accordingly, the output equations (55) and (57) together with equation (61) lead to the set of equations

$$\mathbf{L}_1\mathbf{T} + \mathbf{L}_2\mathbf{C} = 0 \tag{68}$$

$$\mathbf{L}_2\mathbf{F} = 0 \tag{69}$$

$$\mathbf{L}_2\mathbf{G} \neq 0. \tag{70}$$

The solution of these equations provides the solution of the problem, which can again be found in Patton *et al.* (1989) using either the Kronecker canonical form or eigenvector design. Note that this approach is superior to the detection filter approach in that it accounts for modelling errors also, and to the parity space approach in that it also takes into account the sensitivity to the faults.

In many practical situations, the physical conditions to satisfy (58)–(60) for the existence of a UIO do not exist. In this case, one has to find an optimal approximation due to a performance criterion that takes into account the sensitivity with respect to the disturbance $\mathbf{d}$, as well as the sensitivity with respect to faults $\mathbf{f}$. This task can either be solved in the time domain (Wünnenberg and Frank, 1987) or in the frequency domain (Ding and Frank, 1989).

5.5. *Optimal time domain approximation.* For any time k, equations (54) and (55) can be re-written as

$$\begin{bmatrix} \mathbf{y}_{k-s} \\ \mathbf{y}_{k-s+1} \\ \vdots \\ \mathbf{y}_k \end{bmatrix} = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^s \end{bmatrix} \mathbf{x}_{k-s} + H_1 \begin{bmatrix} \mathbf{u}_{k-s} \\ \mathbf{u}_{k-s+1} \\ \vdots \\ \mathbf{u}_k \end{bmatrix}$$

$$+ \mathbf{H}_2 \begin{bmatrix} \mathbf{d}_{k-s} \\ \mathbf{d}_{k-s+1} \\ \vdots \\ \mathbf{d}_k \end{bmatrix} + \mathbf{H}_3 \begin{bmatrix} \mathbf{f}_{k-s} \\ \mathbf{f}_{k-s+1} \\ \vdots \\ \mathbf{f}_k \end{bmatrix} \quad (71)$$

where s is the considered time horizon, and the matrices $\mathbf{H}_1$, $\mathbf{H}_2$ and $\mathbf{H}_3$ are given by

$$\mathbf{H}_1 = \begin{bmatrix} \mathbf{0} & & & & \\ \mathbf{CB} & \mathbf{0} & & \mathbf{0} & \\ \mathbf{CAB} & \mathbf{CB} & \mathbf{0} & & \\ \vdots & & \ddots & \ddots & \\ \mathbf{CA}^{s-1}\mathbf{B} & \cdots & \cdots & \mathbf{CB} & \mathbf{0} \end{bmatrix}, \quad (72)$$

$$\mathbf{H}_2 = \begin{bmatrix} \mathbf{F} & \mathbf{0} & & & \\ \mathbf{CE} & \mathbf{F} & \mathbf{0} & \mathbf{0} & \\ \mathbf{CAE} & \mathbf{CE} & \mathbf{F} & \mathbf{0} & \\ \vdots & & \ddots & \ddots & \mathbf{0} \\ \mathbf{CA}^{s-1}\mathbf{E} & \cdots & \cdots & \mathbf{CE} & \mathbf{F} \end{bmatrix}, \quad (73)$$

$$\mathbf{H}_3 = \begin{bmatrix} \mathbf{G} & \mathbf{0} & & & \\ \mathbf{CK} & \mathbf{G} & \mathbf{0} & \mathbf{0} & \\ \mathbf{CAK} & \mathbf{CK} & \mathbf{G} & \mathbf{0} & \\ \vdots & & \ddots & \ddots & \mathbf{0} \\ \mathbf{CA}^{s-1}\mathbf{K} & \cdots & \cdots & \mathbf{CK} & \mathbf{G} \end{bmatrix}. \quad (74)$$

In order to generate a scalar residual r_k one has to check if the above state equations hold for the available input and output data. This can be done by calculating equation (75) on-line at each sampling time k:

$$r_k = \mathbf{v}^{\mathrm{T}} \left[\begin{bmatrix} \mathbf{y}_{k-s} \\ \mathbf{y}_{k-s+1} \\ \vdots \\ \mathbf{y}_k \end{bmatrix} - \mathbf{H}_1 \begin{bmatrix} \mathbf{u}_{k-s} \\ \mathbf{u}_{k-s+1} \\ \vdots \\ \mathbf{u}_k \end{bmatrix} \right]. \quad (75)$$

The vector $\mathbf{v}^{\mathrm{T}}$ that has to be found is called the residual generator and has to meet the following requirements. In order to make the residual r_k independent of any initial condition $\mathbf{x}_{k-s}$, the vector $\mathbf{v}^{\mathrm{T}}$ has to be chosen such that

$$\mathbf{v}^{\mathrm{T}} \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^s \end{bmatrix} = \mathbf{0}. \quad (76)$$

Furthermore, because the residual has to be affected by the fault, the condition

$$\mathbf{v}^{\mathrm{T}}\mathbf{H}_3 \neq \mathbf{0} \quad (77)$$

must hold. Notice that if $\mathbf{v}^{\mathrm{T}}$ could be determined as to satisfy the additional condition

$$\mathbf{v}^{\mathrm{T}}\mathbf{H}_2 = \mathbf{0} \quad (78)$$

which implies that the residual is *not* affected by the unknown input vector $\mathbf{d}$, then we would arrive at the "ideal" solution.

Equations (76)–(78) can be interpreted as a reformulation of the requirements given by equations (58)–(60). The determination of $\mathbf{v}^{\mathrm{T}}$ provides the solution of the problem.

Lou *et al.* (1986) call the space spanned by the solutions of equation (76) the parity space of order s. The choice of s clearly influences the solvability of equations (76)–(78). For the "ideal" solution an upper bound is generally given by

$$s \leq n, \quad (79)$$

where n is the dimension of $\mathbf{x}_k$. For the special case where the measurements are not disturbed ($\mathbf{F} = \mathbf{0}$), we have (see Patton *et al.*, 1989):

$$s \leq n - m + 1 \quad (80)$$

with m the number of independent measurements. Note that increasing the order s beyond these bounds does not provide additional freedom in the design of the "ideal" filter. However, this does not apply to the approximate solution.

In order to find the optimal approximation, a certain performance index that contains a measure of the effects of the unknown input vector, $\mathbf{d}_k$, and the fault vector, $\mathbf{f}_k$, must be defined. Because no assumptions or restrictions are made about the evolution of the disturbances and the faults, the performance index can only take into account the distribution of these signals into the system. As a proper choice of the performance index we define

$$P = \frac{\|\mathbf{v}^{\mathrm{T}}\mathbf{H}_2\|}{\|\mathbf{v}^{\mathrm{T}}\mathbf{H}_3\|} \quad (81)$$

which must be minimized in connection with equation (76). The symbol $\|\cdot\|$ denotes the Euclidean vector norm. To ensure that $\mathbf{v}^{\mathrm{T}}$ also fulfils equation (76) we reformulate the problem as follows. Find a vector $\mathbf{w}^{\mathrm{T}}$ such that the performance index

$$P = \frac{\|\mathbf{w}^{\mathrm{T}}\mathbf{V}_0\mathbf{H}_2\|}{\|\mathbf{w}^{\mathrm{T}}\mathbf{V}_0\mathbf{H}_3\|} \quad (82)$$

becomes minimal. Here, $\mathbf{V}_0$ has the meaning of a base for all possible solutions to equation (76).

Hence $\mathbf{w}^{\mathrm{T}}$ singles out the best vector $\mathbf{v}^{\mathrm{T}}$ of all possible solutions represented by $\mathbf{V}_0$. The satisfaction of equation (76) guarantees that the residual is not affected by the initial condition $\mathbf{x}_{k-s}$ and neither by the input signal $\mathbf{u}_k$.

Clearly, the choice of the filter degree s has an influence on the "optimal" solution. The least square character of the index shows that an increase of s can actually reduce the value of the performance index.

An appropriate value for s can be found by the designer by a systematic increase of s. Note that the performance index used here is different from the one used in Lou *et al.* (1986) which does not guarantee independence of the residual of each system input $\mathbf{u}_k$.

By inspection of equations (75) and (76) it becomes clear that the residual generator is an observer, as is explicitly shown in Patton *et al.* (1989). Characteristic of this solution is that all the poles of the observer described by equation (75) are equal to zero. The residual generator according to equation (75) is therefore a so-called dead-beat observer, and the residual can be interpreted as its output estimation error. It can also be shown (Patton *et al.*, 1989) that the residual of any observer with arbitrary poles can be generated by low-class filtering of the residual of a dead-beat observer. An according extension of the filter design by using weighting factors in the performance index, P, equation (82), is described in Wünnenberg and Frank (1990).

The solution to the optimization problem according to equation (81) can be obtained by a differentiation of the performance index w.r.t. $\mathbf{w}^T$. This leads to the relation

$$\mathbf{w}^T(\mathbf{V}_0\mathbf{H}_2\mathbf{H}_2^T\mathbf{V}_0^T - P\mathbf{V}_0\mathbf{H}_3\mathbf{H}_3^T\mathbf{V}_0^T) = \mathbf{0} \qquad (83)$$

the solution of which gives the solution of the design problem. This shows that the problem reduces to a generalized eigenvalue–eigenvector problem. The minimal eigenvalue is the optimal value of the performance index and the corresponding eigenvector is the selector for the optimal residual generator $\mathbf{v}^T$.

5.6. *Optimal frequency domain approximation.* Consider the IFD scheme as illustrated in Fig. 10 where the observer is again given by equations (54) and (55). The objective is to optimize the observer such that the residual r becomes robust with respect to the unknown input, $\mathbf{d}$, and sensitive with respect to the faults, $\mathbf{f}$. In these terms the desired properties specified by equations (58)–(60) can be also expressed as follows;

$$I = \sup_{\mathbf{h}^T(j\omega)} I_s = \sup \frac{\left\|\frac{\partial r}{\partial \mathbf{f}}\right\|}{\left\|\frac{\partial r}{\partial \mathbf{d}}\right\|} \qquad (84)$$

$$\lim_{t\to\infty} r = 0 \text{ for all initial conditions } \mathbf{x}(0) \text{ and } \mathbf{z}(0) \qquad (85)$$

where $\mathbf{h}(s)$ represents the transfer function matrix of

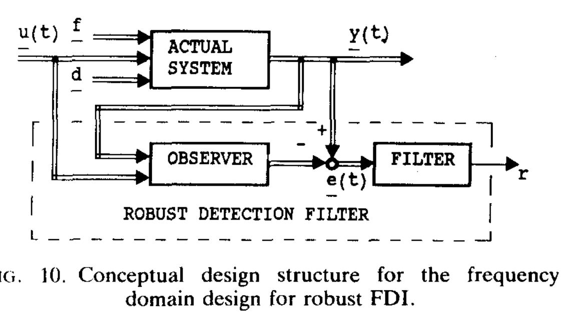

FIG. 10. Conceptual design structure for the frequency domain design for robust FDI.

the observer, and $\|\cdot\|$ denotes the following norm

$$\|\mathbf{q}^T\| = \frac{1}{2\Pi}\int_{-\infty}^{+\infty} \mathbf{q}^T(j\omega)\mathbf{q}(-j\omega)\,d\omega. \qquad (86)$$

This constrained optimization problem can be reduced to the following unconstrained optimization problem:

$$I = \sup_{\mathbf{p}^T(j\omega)} \frac{\|\mathbf{p}^T(j\omega)\mathbf{H}_f(j\omega)\|}{\|\mathbf{p}^T(j\omega)\mathbf{H}_d(j\omega)\|} \qquad (87)$$

with $\mathbf{H}_f(j\omega)$ and $\mathbf{H}_d(j\omega)$ denoting the system transfer function matrices from $\mathbf{f}$ and $\mathbf{d}$, respectively, to $\mathbf{y}$. The desired $\mathbf{p}(s)$ is a real-rational vector which is stable, i.e.

$$\frac{1}{2\Pi}\int_{-\infty}^{+\infty} \mathbf{p}^T(j\omega)\mathbf{p}(-j\omega)\,d\omega < \infty, \qquad (88)$$

and there exists a one-to-one-relation between $\mathbf{p}(s)$ and $\mathbf{h}(s)$. It can be proved that the optimal value of I_s is

$$\sup_{\mathbf{h}^T(j\omega)} I_s = \sup_{\omega}(\bar{\lambda}(\omega)) \qquad (89)$$

where $\lambda(\omega)$ denotes the maximal singular value of the regular pencil

$$\mathbf{H}_f(j\omega)\mathbf{H}_f^T(-j\omega) - \lambda(\omega)\mathbf{H}_d(j\omega)\mathbf{H}_d^T(-j\omega).$$

It can be shown (Ding and Frank, 1989) that the optimal solution $\mathbf{p}_{opt}^T(s)$ consists of two parts and can be written as

$$\mathbf{p}_{opt}^T(s) = \lim_{i\to\infty} g_i(s)\mathbf{v}^T(s) \qquad (90)$$

where $\mathbf{v}^T(s)$ is the eigenvector of the regular pencil, obeying

$$\mathbf{v}^T(j\omega)(\mathbf{H}_f(j\omega)\mathbf{H}_f^T(-j\omega) - \bar{\lambda}(\omega)\mathbf{H}_d(j\omega)\mathbf{H}_d^T(-j\omega)) = 0, \qquad (91)$$

and the scalar function

$$g(s) = \lim_{i\to\infty} g_i(s)$$

describing a frequency selecting filter with order $i \to \infty$ which satisfies for any stable function $\mathbf{W}(s)$:

$$(1)\ g(j\omega)\mathbf{W}^T(j\omega) = \mathbf{0} \quad \omega \neq \omega_0 \qquad (92)$$

$$(2)\ \int_{-\infty}^{+\infty} g(j\omega)\mathbf{W}^T(j\omega)\mathbf{W}(-j\omega)g(-j\omega)\,d\omega = \mathbf{W}^T(j\omega_0)\mathbf{W}(-j\omega_0). \qquad (93)$$

In our case ω_0 is the frequency at which $\bar{\lambda}(\omega)$ reaches its maximal value, namely

$$\bar{\lambda}(\omega_0) = \sup_{\omega}(\bar{\lambda}(\omega)). \qquad (94)$$

In practice one cannot realize an ideal frequency-selecting filter as required by equations (92) and (93). But by the progress of computer technology it has become feasible to design a frequency-selecting filter which approaches this with high accuracy. Thus, instead of the solution in equation (90) one has a sub-optimal solution according to

$$\mathbf{p}^T(s) = g_f(s)\mathbf{v}^T(s). \qquad (95)$$

$g_f(s)$ represents a realizable frequency-selecting filter which can be obtained using digital filter techniques.

In summary, the design of the optimal detection filter consists of the following four steps;

—determination of maximal singular value $\bar{\lambda}(\omega)$ and the corresponding eigenvector $\mathbf{v}^T(j\omega)$;
—determination of the frequency ω_0;
—design of a frequency selecting filter with respect to ω_0;
—determination of $\mathbf{h}^T(s)$.

6. *Optimally robust observer schemes for IFD, CFD and AFD*

In this section we show how to build observer schemes using unknown input observers for the tasks of IFD, CFD and AFD.

6.1. *General structure for IFD, CFD and AFD.*

In general, the number of observers in an observer scheme can be arbitrary. To become specific and for the sake of simplicity of presentation let us assume that m different faults f_i $(i = 1, \ldots, m)$ can occur in the system where m is also the number of measurements available. The general structure of the robust observer scheme—no matter whether being used for instrument, actuator or component fault detection—is shown in Fig. 11.

Concerning the design of the UIOs one has to partition the faults into subsets of $\mathbf{f}$, specified by vectors $\mathbf{f}_i$. Each UIO is now assigned to be sensitive to a different fault or set of faults and invariant to the other faults. The remaining design freedom is then used to provide invariance to unknown inputs. This is done, in turn, to such an extent that properly structured sets of residuals, r_i, are obtained which allow a unique decision of the time of occurrence and location of the faults. The final design of the observers depends on the number of faults that are to be detected at the same time. The three extremes are:

(a) faults have to be detected, but not isolated;
(b) only a single fault is to be detected and isolated at a time;
(c) all faults are to be detected and isolated even if they occur simultaneously.

Case (*a*). The FDI-Scheme reduces to a single UIO that generates a residual that is sensitive to all faults whilst being robust to $m - 1$ unknown inputs. The full design freedom can be used to generate a residual that is robust to the maximum number of unknown inputs.

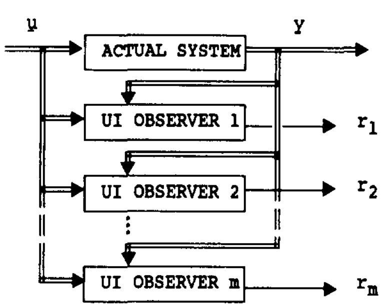

FIG. 11. General structure of the Unknown Input Observer Scheme.

Case (*b*). One can achieve the most robust FDI-scheme that allows fault isolation because the maximum design freedom is left for the generation of robustness to unknown inputs. Here, the ith observer $(i = 1, 2, \ldots, m)$ is designed such that it becomes invariant to the ith fault, f_i, and of $m - 2$ unknown inputs. In other words, f_i is interpreted as an unknown input and the remaining design freedom can be used for generating invariance with respect to $m - 2$ other (genuine) unknown inputs, e.g. for modelling errors. Repeating this design m times one arrives at an UIO scheme according to Fig. 11. Here the first residual, r_1, depends on all but the first fault, the second residual, r_2, on all but the second fault and so on; that is,

$$\begin{aligned} r_1 &= q_1(\mathbf{f}_2, \mathbf{f}_3, \ldots, \mathbf{f}_m) \\ r_2 &= q_2(\mathbf{f}_1, \mathbf{f}_3, \ldots, \mathbf{f}_m) \\ &\vdots \\ r_i &= q_i(\mathbf{f}_1, \ldots, \mathbf{f}_{i-1}, \mathbf{f}_{i+1}, \ldots, \mathbf{f}_m) \\ &\vdots \\ r_m &= q_m(\mathbf{f}_1, \ldots, \mathbf{f}_{m-1}). \end{aligned} \tag{96}$$

The decision function for the logical evaluation of the residuals could then be as follows:

(a) if r_2 and r_3 and $\cdots$ and r_m are unequal to zero (and $r_1 = \mathbf{0}$), the first fault, $\mathbf{f}_1$, occurs;
(b) if r_1 and r_3 and $\cdots$ and r_m are unequal to zero (and $r_2 = \mathbf{0}$), the second fault, $\mathbf{f}_2$, occurs;

and so forth. It is seen that only a single fault at a time can be detected.

Case (*c*). To be able to detect and isolate all faults occurring simultaneously, one has to interpret all but the ith fault in the ith observer $(i = 1, 2, \ldots, m)$ as unknown inputs. Therefore the rank of the matrix $\mathbf{E}$ is increased by $m - 1$, which is the largest possible rank of $\mathbf{E}$ for which complete invariance can be achieved in many cases arising in practice. In this case, however, the observer scheme cannot be made robust with respect to any unknown input since no design freedom is left. The residuals depend on the faults due to the following relations:

$$\begin{aligned} r_1 &= q_1'(\mathbf{f}_1) \\ r_2 &= q_2'(\mathbf{f}_2) \\ &\vdots \\ r_i &= q_i'(\mathbf{f}_i) \\ &\vdots \\ r_m &= q_m'(\mathbf{f}_m). \end{aligned} \tag{97}$$

This permits a unique detection and isolation of m faults even if they occur simultaneously. The price to pay is the loss of robustness with respect to unknown inputs acting on the system. Note that this observer scheme is equivalent to the Fault Detection Filter (Beard, 1971; Jones, 1973).

6.2. *Optimally robust CFD.*

The application of the UIO philosophy to component fault detection in large, complex systems can be improved by exploiting more structural insight in the process. If no assumption on the fault mode can be made, a logical approach is to decompose the process and apply a hierarchical scheme of local observers to the system under consideration (Frank, 1987a).

The difficulty of local state observation lies in the couplings among the components. If the couplings are sufficiently weak or measurable, the local observer scheme can be configured such that a malfunction in any of the components affects only the residual of the corresponding local observer. It is thus possible to identify the faulty component uniquely, and the design freedom can be used to improve the robustness to unknown inputs. The resulting local observer scheme is known as the Available-State-Coupled Observer Scheme (ASCOS) (Frank, 1987a).

If the subsystems interact considerably with each other and the coupling signals are not measurable, one may interpret these couplings as unknown inputs acting on the local subsystem. The design freedom for the local observer can then be used to decouple from these signals, and by this the capability of the detection and isolation of the faulty component can be improved. The resulting observer scheme is known as the Estimated-State-Coupled Observer Scheme (ESCOS) (Frank, 1987a). In large complex systems one has to build a hierarchical structure to supervise the whole system. This hierarchical structure consists of different observer schemes, operating on different system levels.

The most simple structure is that of the local observer scheme as described in the previous section under case (a), indicating that a fault happened in the corresponding subsystem. The next step is a bank of filters according to case (b), indicating that a single fault occurred in this subsystem. The observer bank, designed according to case (c), is able to detect simultaneously appearing faults. Note that by structuring the observers for a single subsystem in this manner, one reaches the highest degree of robustness possible.

In the next level of the hierarchical structure the observer scheme has principally the same configuration. The only difference is that a set of subsystems is interpreted as a new subsystem with the corresponding input and output signals.

This may be repeated until the highest level possible is reached, which is when one designs the observers for the overall system. The resulting structure is known as the Hierarchical Observer Scheme (HOS) (Frank, 1987a).

6.3. *Optimally robust AFD.* The detection of actuator faults can be viewed as a special case of component fault detection and can therefore be achieved by an observer scheme that is equivalent to the structure shown in Fig. 11. Yet it can be simplified because the location of the fault is identical with that of an input signal. For simplification it is assumed that m inputs act onto the system and a single actuator fault may possibly appear in any of the input channels. The structure then obtained is as shown in Fig. 12.

One can see that any observer is driven by all but one input and all outputs of the system. The ith residual is sensitive to all but the ith actuator fault and the detection of a single fault at a time is possible. The ith input that acts on the system must be treated like

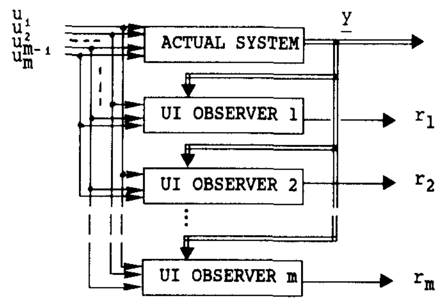

FIG. 12. Generalised Observer Scheme for AFD.

an unknown input and the ith observer must be robust to this missing signal.

If m actuator faults are to be detected at a time, the resulting structure of the observer scheme differs from that of Fig. 12 in that the ith observer is only driven by the ith input ($i = 1, 2, \ldots, m$). The ith residual is then only affected by the ith actuator fault. The other $m - 1$ input signals must be interpreted as unknown inputs, i.e. the residual has to be made robust to these missing signals.

6.4. *Optimally robust IFD.* The application to the special case of instrument fault detection allows more transparency in the FDI structures, similar to AFD, because the faults are associated with direct changes in the corresponding output signals. Instrument faults are therefore easier to detect than component faults.

Suppose a single fault at a time in one of the m sensors of a system is to be detected with robustness to as many unknown inputs as possible. Then one arrives at the so-called *generalized observer scheme* (GOS) for IFD (Frank, 1987a). This is illustrated in Fig. 13.

The ith observer ($i = 1, 2, \ldots, m$) is driven by all but the ith measured variable (i.e. $y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_m$), and is made invariant to $m - 2$ unknown inputs. y_i is not used in the ith observer because y_i is assumed to be corrupted by the fault and therefore does not carry information about the system.

When, on the other hand, the system can be considered undisturbed, and hence no robustness is required, but simultaneous faults in all m instruments are to be detected one arrives at the Dedicated Observer Scheme (DOS) as proposed by Clark

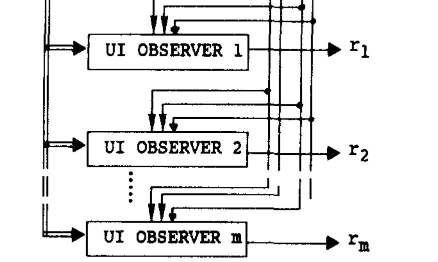

FIG. 13. Generalised Observer Scheme (GOS) for IFD.

(1978b), where the ith observer is driven by the ith output only. Here, the ith fault only affects the ith residual, r_i, which can easily be evaluated in a decision logic. If the system is observable from any of the outputs, each observer is a full order Luenberger observer estimating the full state, and the ith residual is the output estimation error $y_i - \hat{y}_i$.

7. *Fault diagnosis using knowledge-based models*

Knowledge-based methods (expert systems) complement the existing analytical and algorithmical methods of fault detection: they open a new dimension of possible fault diagnosis for complex processes with incomplete process knowledge. Whilst the algorithmic methods use *quantitative* analytical models, the expert system approach makes use of *qualitative* models based on the available knowledge of the system. The combination of both strategies allows the evaluation of all available information and knowledge of the system for fault detection. Such additional knowledge may be, for example, the degree of ageing, the operational environment, used tools, history of operation fault statistics etc.

The architecture of a fault diagnosis system using knowledge-based models is shown in Fig. 14. The core is an on-line expert system which complements the analytical model-based method with the knowledge-based method using heuristic reasoning. The resulting overall fault detection system consists of the following architectural components;

—the knowledge base (knowledge of facts and rules)

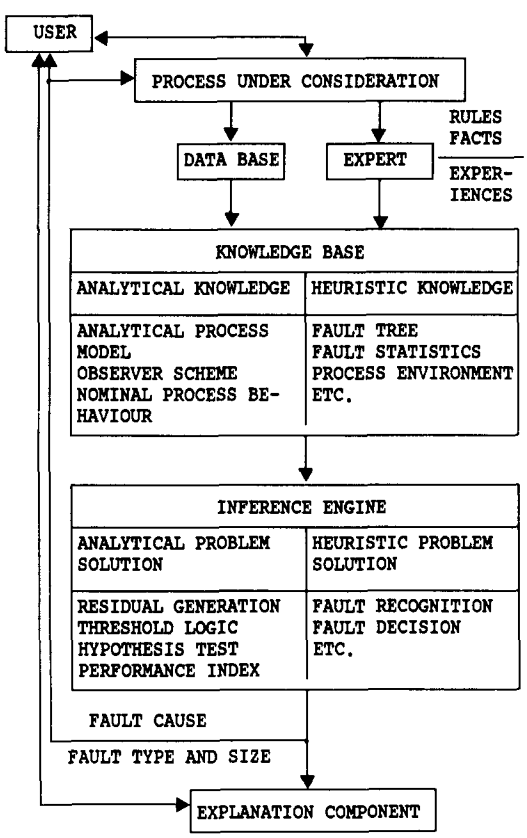

FIG. 14. Architecture of a model- and knowledge-based fault diagnosis system.

—the data base (information about the present state of the process)

—the inference engine (forward or backward reasoning)

—the explanation component (to inform the user on why and how the conclusions were drawn).

Knowledge of the system is acquired from an expert who may be the same person as the user.

As can be seen from Fig. 14, the resulting on-line expert system combines the analytical redundancy FDI method as described earlier in this paper with the method of fault diagnosis by evaluation of heuristic knowledge about the process. This task is done in the inference engine which has to combine heuristic reasoning with algorithmic operations in terms of the evaluation of analytical redundancy. The inference engine has access to

—the analytical knowledge in terms of the mathematical model (structure and parameters)

—heuristic knowledge of fault propagation, fault statistics, operational and environmental conditions, process history etc.

—the actual data (inputs, outputs, operating conditions etc.).

For more details we refer to the chapter of S. Tzafestas in the book by Patton *et al.* (1989).

8. *Conclusion*

In this paper we have discussed the analytical redundancy approach to FDI in dynamic systems. It has been pointed out that there exist various techniques and very elaborate procedures ready for application. Simulation studies and experimental results have shown that the FDI schemes using analytical redundancy have reached a certain degree of maturity. There are, in particular, a number of encouraging results in the application to mechanical systems as, for example, aircraft or advanced transport systems. It should be noted, however, that in cases where only poor or imprecise analytical models are available, as, for example, in chemical plants, the model-based FDI approach is still problematical. In such cases the support by knowledge-based methods may be unavoidable.

In conclusion, the question of application of any model-based FDI scheme is primarily a question of the quality of the available mathematical model of the system. In addition to this, the reachable quality of fault isolation decisively depends on the number of available measurements.

Acknowledgements—The author wants to thank Mr J. Wünnenberg and Mr Ding for esssential scientific contributions.

References

Basseville, M. and A. Benveniste (Eds) (1985). Detection of Abrupt Changes in Signals and Dynamical Systems. *Lecture Notes in Control and Information Sciences*, Vol. 77. Springer, Berlin.

Beard, R. V. (1971). *Failure Accommodation in Linear Systems Through Self-Reorganization*. Dept. MVT-71-1, Man Vehicle Laboratory, Cambridge, MA.

Chow, E. Y. and A. S. Willsky (1984). Analytical

redundancy and the design of robust failure detection systems. *IEEE Trans. Aut. Control,* **AC-29,** 603–614.
Clark, R. N. (1978a). A simplified instrument failure detection scheme. *IEEE Trans. Aerospace Electron. Syst.,* **14,** 558–563.
Clark, R. N. (1978b). Instrument fault detection. *IEEE Trans. Aerospace Electron. Syst.,* **14,** 456–465.
Clark, R. N., D. C. Fosth and V. M. Walton (1975). Detection instrument malfunctions in control systems. *IEEE Trans. Aerospace Electron. Syst.,* **AES-11,** 465–473.
Deckert, J. C., M. N. Desai, J. J. Deyst and A. S. Willsky (1977). DFBW sensor failure identification using analytic redundancy. *IEEE Trans. Aut. Conrol,* **AC-22,** 795–809.
Desai, M. and A. Ray (1981). A fault detection and isolation methodology. *Proc. 20th Conf. on Decision and Control,* 1363–1369.
Ding, X. and P. M. Frank, (1989). Komponentenfehlerdetektion mittels auf Empfindlichkeitsanalyse basierender robuster Detektionsfilter *at Automatisierungstechnik.* Oldenburg, Munich. (To be published).
Emami-Naeini A. *et al.* (1986). *Robust detection, isolation, and accommodation for sensor failures.* NASA Contractor Report No. CR-174825.
Frank, P. M. (1987a). Fault diagnosis in dynamic systems via state estimation—A Survey. In Tzafestas, S., M. Singh and G. Schmidt (Eds), *System Fault Diagnostics, Reliability and Related Knowledge-Based Approaches,* Vol. 1, pp. 35–98. Reidel, Dordrecht.
Frank, P. M. (1987b). Advanced fault detection and isolation schemes using nonlinear and robust observers. Presented at 10th IFAC World Congress, Munich, 27-31 July 1987.
Frank, P. M. (1988). Fault diagnosis on the basis of dynamic process models. Presented at 12th IMACS World Congress on Scientific Computation, Paris, 18–22 July 1988.
Frank, P. M. and L. Keller (1980). Sensitivity discriminating observer design for instrument failure detection. *IEEE Trans. Aerospace Electron. Syst.,* **AES-16,** 460–467.
Frank, P. M. and L. Keller (1984). Entdeckung von Instrumentenfehlanzeigen mittels Zustandsschätzung in technischen Regelungssystemen. *Fortschrittsbericht VDI-Z,* Reihe 8, Nr. 80. VDI, Düsseldorf.
Ge, W. and C. Z. Fang (1988). Detection of faulty components via robust observation. *Int. J. Control,* **47,** 581–599.
Himmelblau, D. M. (1978). *Fault Detection and Diagnosis in Chemical and Petrochemical Processes.* Elsevier, NY.
Isermann, R. (1984). Process fault detection based on modeling and estimation methods—a survey. *Automatica,* **20,** 387–404.
Jones, H. L. (1973). Failure detection in linear systems. Ph.D. Thesis, MIT, Cambridge, MA.
Kitamura, M. (1980). Detection of sensor failures in nuclear plant using analytic redundancy. *Trans. Am. Nucl. Soc.,* **34,** 581–583.
Labarrère, M. (1987). Aircraft Sensor Failure Detection by Analytic Redundancy. In Singh, M. G. (Ed.), *Systems and Control Encyclopedia,* Vol. 1, pp. 246–251. Pergamon Press, Oxford.
Lou, X. C., A. S. Willsky and G. L. Verghese (1986). Optimally robust redundancy relations for failure detection in uncertain systems. *Automatica,* **22,** 333–344.
Massoumnia, M. A. (1986). A geometric approach to failure detection and identification in linear systems. Ph.D. Thesis, MIT, Cambridge, MA.
Mehra, R. K. and I. Peshon (1971). An innovations approach to fault detection and diagnosis in dynamic systems. *Autmatica,* **7,** 637–640.
Milne, R. (1987). Strategies for diagnosis. *IEEE Trans. Syst., Man Cybern.,* **SMC-17,** 333–339.
Montgomery, R. C. and A. K. Caglayan (1974). A self-reorganizing digital flight control system for aircraft. Presented at AIAA 12th Aerospace Sciences Meeting, Washington, D.C., 30 Jan.-1 Feb. 1974.
Onken, R. and N. Stuckenberg (1979). Failure detection in signal processing and sensing in flight control systems. *Proc. IEEE Conf. on Decision and Control,* San Diego, pp. 449–454.
Patton, R. J., P. M. Frank and R. N. Clark (Ed) (1989). *Fault Diagnosis in Dynamic Systems, Theory and Applications.* Prentice-Hall, Englewood Cliffs, NJ.
Patton, R. J., S. W. Willcox and J. S. Winter (1987). A parameter insensitive technique for aircraft sensor fault analysis. *AIAA J. Guidance Control Dynam.,* 359–367, July–August 1987.
Potter, I. E. and M. C. Sunman (1977). Thresholdless redundancy management with arrays of skewed instruments. *Integrity in Electronic Flight Control Systems,* AGARDOGRAPH-224, 15-11 to 15-25.
Viswanadham, N. and R. Srichander (1987). Fault detection using unknown-input observers. *Control Theory and Advanced Technology,* Vol. 3, pp. 91–101. MITA.
Viswanadham, N., V. V. S. Sarma and M. G. Singh (1987). *Reliability of Computer and Control Systems,* North Holland Systems and Control Series, Vol. 8. North Holland, Amsterdam.
Walker, B. K. (1983). Recent Developments in Fault Diagnosis and Accommodation. Presented at AIAA Guidance and Control Conf., Gatlinburg, TN.
Watanabe, K. and D. M. Himmelblau (1982). Instrument fault detection in systems with uncertainties. *Int. J. Syst. Sci.,* **13,** 137–158.
Weiss, J. L., K. R. Pattipati, A. S. Willsky, J. S. Eterno and J. T. Crawford (1985). Robust detection/isolation/accommodation for sensor failures. *NASA Contr. Rep. 174797, Lewis Res. Cent. NAS* 3-24078, Alphatech Inc., Burlington, MA.
Wilbers, D. N. and J. L. Speyer (1989). Detection filters for aircraft sensor and actuator faults. *Proc. ICCON '89 Int. Conf. on Control and Applications,* Jerusalem, April 1989.
Willsky, A. S. (1976). A survey of design methods for failure detection in dynamic systems. *Automatica,* **12,** 601–611.
Wünnenberg, J. (1990). Linear and nonlinear robust fault detection in dynamic systems. Dissertation, University of Duisburg, Fachgebiet Meß- und Regelungstechnik, FRG.
Wünnenberg, J. and P. M. Frank (1987). Sensor Fault Detection via Robust Observers. In Tzafestas, S., M. Singh and G. Schmidt (Eds), *System Fault Diagnostics, Reliability and Related Knowledge-Based Approaches* Vol. 1, pp. 147–160. Reidel, Dordrecht.
Wünnenberg, J. and P. M. Frank, (1990). Dynamic model-based incipient fault detection concept for robots. Submitted to the 11th IFAC World Congress, 13–17 August 1990, Tallinn.