Appl. Math. and Comp. Sci., 1991, vol.1, No.1, 83–106

FAULT DETECTION IN DYNAMIC SYSTEMS. STATE ESTIMATION APPROACH

JÓZEF KORBICZ*, ZOHREH FATHI**, W. FRED RAMIREZ**

The state estimation problem for dynamic systems is one of the fundamental problems in the fields of modeling, optimal control, and fault detection and diagnosis. The linear and nonlinear state estimation has been a very active research field during the last 30 years. The purpose of this paper is to give a brief review of the basic fault detection and diagnosis methods based upon the analytical and knowledge–based redundancy. The main emphasis is placed upon estimation methods that are widely applied for fault detection. The advantages and disadvantages of these methods also are discussed both in general and in diagnostic applications.

1. Introduction

Contemporary automatic control systems are becoming more complex to achieve the demanded characteristics of technological processes. Control algorithms are also getting to be more sophisticated. Consequently, there is a growing demand for fault–tolerance which is traditionally achieved through the use of hardware redundancy. In such an approach, the repeated hardware elements (actuators, measurement sensors, process components, etc.) are usually distributed spatially around the system to provide protection against localized damage. Usually a triplex or quadruplex redundancy configuration is applied and redundant outputs are compared for consistency. This approach to fault–tolerance is simple from a theoretical point of view and in many cases is reasonably straightforward to apply, and thus it is widely used in cost–is–no–object situations such as aerospace vehicles or nuclear power related systems.

* Department of Applied Mathematics and Computer Science, Higher College of Engineering, Zielona Góra, Poland

** Department of Chemical Engineering, University of Colorado, Boulder, Colorado U.S.A.

In most actual plants, the major problems associated with hardware redundancy are the extra cost and software. To overcome these and other problems, another approach to fault detection and isolation in automatic processes using analytical (physical or artificial) redundancy has been proposed since the early 1970's. (Himmelblau, 1978; Patton *et al.*, 1989; Thompson and Fleming, 1990). Hardware and analytical redundancies are distinguished in the following manner. For example, the two sensors, s_1 and s_2, are hardware redundant if both measure an identical variable (noise and other characteristics of s_1 and s_2 may not be the same), and analytically redundant if they measure different variables but the variable that s_1 measures can be observed from s_2 and vice versa.

Functionally–redundant fault detection and isolation (FDI) schemes are basically signal processing techniques employing state estimation, parameter estimation, adaptive filtering, variable threshold logic, and statistical decision theory. In general, the FDI schemes use mathematical process models, state estimation techniques, and statistical decision algorithms. It is reasonable to point out that the appeal of the analytical redundancy is in the fact that it can be simply evaluated using the information under well–featured operational conditions without the need of additional physical instrumentation in the plant.

In general, the problem of fault detection in dynamic systems can be formulated as the detection problem of changes of unknown magnitudes which occur at unknown times in controllers, sensors, and/or components of the actual plan. Fault detection and diagnosis for actual plants is important in two aspects. The first aspect is an improvement in system availability, and the second, which is more important, is the protection from disasters. From a technical point of view, it seems that the realizability of a practical detection and diagnosis system has been fairly enhanced by the remarkable development of computer processing technology.

To date many fault detection and diagnosis methods have been proposed for stochastic and deterministic dynamical systems. There exist several surveys of these approaches. Willsky (1976) is a classic paper, discussing many of the techniques available at the time. More recently, an excellent survey was published by Frank (1990) where he presents more important techniques of model–based residual generation using state and parameter estimation methods with emphasis on the latest attempts to achieve robustness with respect to modeling errors. Other earlier surveys have been prepared by Isermann (1984), Himmelblau (1986), and Besseville (1988).

The interested reader can refer to Himmelblau's book (1978) for the fault detection and diagnosis in chemical processes, and to the book edited by Basseville and Benveniste (1986) and the recently published book edited by Patton *et al.* (1989) for the fault detection and isolation techniques based on the use of mathematical models of process systems.

In general, various known approaches to FDI problems using analytical redundancy can be traced back to a few basic concepts. Among these are:

- detection filter (Jones, 1973),
- innovation test using Kalman filters or Luenberger observers (Mehra and Peschon, 1971; Clark *et al.*, 1975; Yoshimura *et al.*, 1979; Kerr, 1982; Watanabe and Himmelblau, 1982; Loparo *et al.*, 1991),
- parity space approach (Deckert *et al.*, 1977; Gertler and Singer, 1990; Gertler *et al.*, 1990; Luck and Ray, 1991),
- parameter estimation technique (Kitamura, 1980; Iserman, 1984),
- expert system applications (Tzafestas, 1989; see Tzafestas in the book by Patton *et al.*, 1989; Neumann, 1990),
- neural networks applications (Naidu *et al.*, 1990; Yao and Zafiriou, 1990; Haesloop and Holt, 1990).

Among the above mentioned methods and techniques are the expert system and neural networks approaches, which are especially interesting and important from a practical point of view. They complement the existing analytical and algorithmic methods of fault detection by application of artificial intelligence (Miller *et al.*, 1990; Johannsen and Alty, 1991). The main advantage of the expert system approach lies in the fact that it makes use of qualitative models, based on the available knowledge of the system, and quantitative analytical models. The combination of both strategies allows the use of all available information given by numeric and symbolic models for performing the fault detection and diagnosis task. In general, the architecture of a fault detection and diagnosis system using knowledge- -based models is shown in Figure 1.

We should point out that the latest advances in the area of artificial intelligence, in particular in the neural networks (Miller *et al.*, 1990), provide the potential for new approaches to fault detection and diagnosis in dynamic systems. Such trainable networks have been successfully used in sensor failure detection for control systems by Naidu *et al.* (1990) and Yao and Zafiriou (1990). Application of neural networks to failure state recognition in chemical plants has been studied by Venkatasubramanian and Chen (1989) and Watanabe *et al.* (1989). At the present state of development, it seems

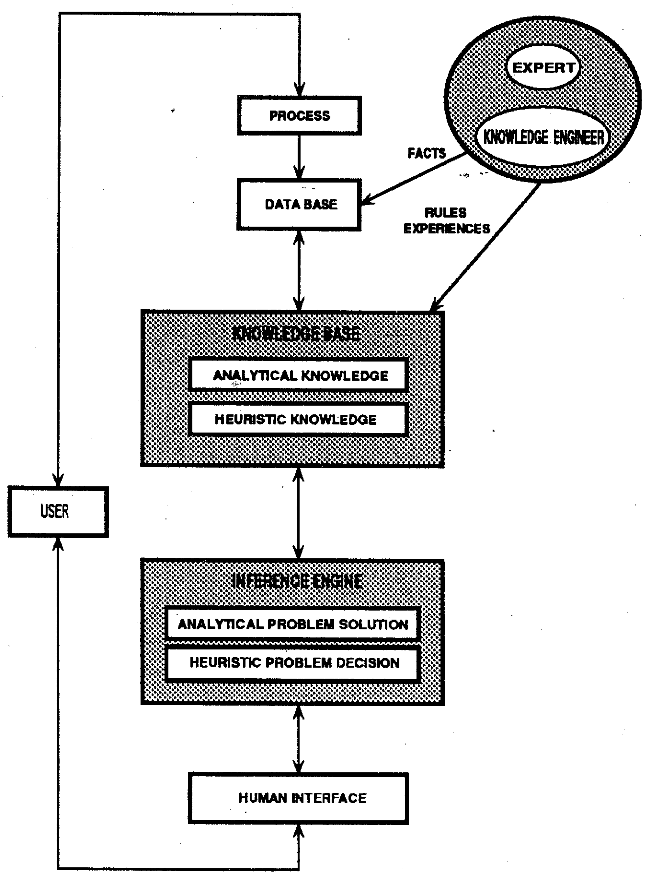

Figure 1. The fault detection and diagnosis system architecture

that the main advantage of the use of neural networks is the improvement of the failure cognition system after on-line implementation. Such designed systems can be trained on-line to improve the initial off-line training by learning to avoid false alarms from which it may initially suffer.
The extended Kalman filter considered below has been applied to the learning algorithms of layered neural networks by Watanabe *et al.* (1991).

As mentioned previously, the research in the area of fault detection and diagnosis in dynamic systems using analytical and knowledge–based redundancy has been conducted by many authors. Most of the work that considers this problem from the point of analytical redundancy proposes to use the Kalman filter in the stochastic case, and the Luenberger observer in the deterministic case. For example, Mehra and Peschon (1971) have shown how the innovation properties can be used for fault detection (tests of whiteness, mean, covariance, and chi–square). The detector of Yoshimura *et al.* (1979) that considers parametric failures is composed of a normal mode filter and an adaptive extended Kalman filter. The latter filter estimates the system parameters and the state under the failure mode. Another strategy (a two–level approach, a Kalman filter for the states and a least squares estimator for the parameters) has been proposed by Watanabe and Himmelblau (1983) for process models nonlinear in states but linear in coefficients. An adaptive filtering–based method for failure detection has been considered by Willsky *et al.* (1974). Halme and Selkainaho (1986) introduced the idea of partitioning the system into smaller submodels and applied an extended Kalman filter to detect sensor faults. Few works have been completed to date on the study of FDI schemes using nonlinear estimators (Misawa and Hedrick, 1988) in case of nonlinearities in the process. The nonlinear Luenberger state observers approaches of Hengy and Frank (1986) and Wunnenberg and Frank (1990) are the nonlinear local state observers for component fault detection. A nonlinear filtering approach based on a reparameterization of the Kalman filter has been proposed by Davis (1975) and Loparo *et al.* (1986) for failure detection problems. Some applications of nonlinear filtering approach to process diagnosis and failure detection are presented by Loparo *et al.* (1991) for leak detection in a heat exchanger process and Eckert *et al.* (1986) for instrument failure in a pressurized water reactor.

Each of the cited references represents a different approach to fault detection and diagnosis, but most apply the Kalman filter algorithms in different ways to solve the problem. State estimation approaches have been a very active research field in the last 30 years both from a theoretical (Anderson and Moore, 1979) and applications (Sorenson, 1985; Ramirez, 1987) point of view. In this paper a survey of currently available recursive state estimation techniques applicable to a broad class of linear and nonlinear systems is presented. Especially, the focus will be on those Kalman filter algorithms which are actually applied or can be applied for fault detection and diagnosis systems.

The methods that are considered in this paper are the linear Kalman filter, the extended Kalman filter, the adaptive filtering approach, the nonlinear Kalman filter, the nonlinear approach for systems with coupled static and dynamic models, and the suboptimal filtering algorithms for linear and nonlinear systems. The advantages and disadvantages of these methods are discussed both in general and in the area of diagnostic applications.

2. Linear Kalman Filter

In 1960 Kalman introduced the concept of an optimal linear filter minimizing the mean square estimation error that is now known as the Kalman filter. The Kalman filter provides estimates of the state vector for a given system at the current time based on the sequence of measurements including the present time.

2.1. Static System

Very often the process under consideration can be represented by a static model, or in other words, by a steady–state vector. Suppose we want to estimate the n–dimensional state vector $\boldsymbol{x}$ of a static system described by

$$\mathbf{0} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{\theta}) + \boldsymbol{w} \tag{1}$$

and observed according to

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{v} \tag{2}$$

where $\boldsymbol{f}(\cdot)$ denotes an n–dimensional nonlinear vector function, $\boldsymbol{u}$ is a p–dimensional input (control) vector, $\boldsymbol{\theta}$ is an s–dimensional parameter vector, $\boldsymbol{y}$ is an m–dimensional output measurement vector $(n \geq m)$, and $\boldsymbol{H}$ is a known measurement matrix $(m \times n)$.

It is assumed that the system uncertainty $\boldsymbol{w}$ and the measurement errors $\boldsymbol{v}$ have the following characteristics

$$\begin{aligned} E[\boldsymbol{w}] &= \mathbf{0} \qquad & E[\boldsymbol{v}] &= \mathbf{0} \\ E[\boldsymbol{w}\,\boldsymbol{w}^T] &= \boldsymbol{Q} \qquad & E[\boldsymbol{v}\,\boldsymbol{v}^T] &= \boldsymbol{R} \end{aligned} \tag{3}$$

where $\boldsymbol{Q}(n \times n)$ and $\boldsymbol{R}(m \times m)$ are known covariance matrices and $E[\cdot]$ denotes the expectation operator.

On minimizing a weighted–least–squares performance index (Bryson and Ho, 1975), the solution for the optimal estimate of $\boldsymbol{x} = \widehat{\boldsymbol{x}}$ is as follows

$$\overline{\boldsymbol{x}} = \mathrm{Sol}_x\{\boldsymbol{f}(\overline{\boldsymbol{x}}, \boldsymbol{u}, \boldsymbol{\theta}) = \boldsymbol{0}\} \tag{4}$$

$$\widehat{\boldsymbol{x}}(i) = \overline{\boldsymbol{x}} + \boldsymbol{P}\boldsymbol{H}^T\boldsymbol{R}^{-1}\left[\boldsymbol{y}(i) - \boldsymbol{H}\overline{\boldsymbol{x}}\right], \quad i = 1, 2, ..., N \tag{5}$$

$$\boldsymbol{P} = \boldsymbol{Q} - \boldsymbol{Q}\boldsymbol{H}^T(\boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^T + \boldsymbol{R})^{-1}\boldsymbol{H}\boldsymbol{Q} \tag{6}$$

where $\mathrm{Sol}_x\{\cdot\}$ denotes the solution of the static equation, $\overline{\boldsymbol{x}}$ is the state estimate before measurements are taken, i denotes the measurement number, and $\boldsymbol{P}$ is the covariance matrix of the error in the state estimate.

It is worthwhile to point out that after the estimation process is completed, the actual value of the performance index J given by

$$J = \frac{1}{2}\left\{(\boldsymbol{x}-\overline{\boldsymbol{x}})^T\boldsymbol{Q}^{-1}(\boldsymbol{x}-\overline{\boldsymbol{x}}) + (\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x})^T\boldsymbol{R}^{-1}(\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x})\right\}; \quad \boldsymbol{x} = \overline{\boldsymbol{x}} + \boldsymbol{w} \tag{7}$$

should be close to $(n+m)/2$, its prior expected value.

2.2. Dynamic Models

Consider a class of discrete–time, linear stochastic systems described by the following equation

$$\boldsymbol{x}(k+1) = \boldsymbol{A}(k)\boldsymbol{x}(k) + \boldsymbol{B}(k)\boldsymbol{u}(k) + \boldsymbol{C}(k)\boldsymbol{w}(k) \tag{8}$$

where k is a discrete time, $\boldsymbol{x}(k)$ is an $n-$dimensional state vector, $\boldsymbol{u}(k)$ is a $p-$dimensional input (control) vector, $\boldsymbol{w}(k)$ is a $q-$dimensional system noise vector, $\boldsymbol{A}(k)$ is a known $(n \times n)-$dimensional system matrix, $\boldsymbol{B}(k)$ is a known $(n \times p)-$dimensional input matrix, and $\boldsymbol{C}(k)$ is a known $(n \times q)-$dimensional system noise matrix.

Assume that the measurement system for the process (8) is given by the algebraic equation

$$\boldsymbol{y}(k) = \boldsymbol{H}(k)\boldsymbol{x}(k) + \boldsymbol{v}(k) \tag{9}$$

where $\boldsymbol{y}(k)$ is an $m-$dimensional output measurement vector, $\boldsymbol{v}(k)$ is an $m-$dimensional observation noise vector, and $\boldsymbol{H}(k)$ is a known $(m \times n)-$dimensional observation matrix. Note that $dim\boldsymbol{x} \geq dim\boldsymbol{y}$.

It is assumed that the system noise $\boldsymbol{w}(k)$, the measurement noise $\boldsymbol{v}(k)$, and the initial condition $\boldsymbol{x}(0)$ are Gaussian random variables with known statistics

$$\begin{array}{lll} E[\boldsymbol{w}(k)] = \mathbf{0} & E[\boldsymbol{v}(k)] = \mathbf{0} & E[\boldsymbol{x}(0)] = \hat{\boldsymbol{x}}(0|0) \\ E[\boldsymbol{w}(k)\boldsymbol{w}^T(l)] = \boldsymbol{Q}(k)\delta_{kl} & E[\boldsymbol{v}(k)\boldsymbol{v}^T(l)] = \boldsymbol{R}(k)\delta_{kl} & E[\boldsymbol{x}(0)\boldsymbol{x}^T(0)] = \boldsymbol{P}(0|0) \end{array} \tag{10}$$

where $\boldsymbol{Q}(k)$ is a $(q \times q)$ symmetric non–negative definite matrix, $\boldsymbol{R}(k)$ is an $(m \times m)$ symetric positive–definite matrix, $\boldsymbol{P}(0|0)$ is an $(n \times n)$ symmetric non–negative definite matrix, and δ_{kl} denotes the Kronecker delta function.

In addition, it is assumed that the random sequences $\boldsymbol{w}(k)$, $\boldsymbol{v}(k)$, and the random variable $\boldsymbol{x}(0)$ are uncorrelated,

$$E[\boldsymbol{w}(k)\boldsymbol{x}^T(0)] = \mathbf{0} \quad E[\boldsymbol{v}(k)\boldsymbol{x}^T(0)] = \mathbf{0} \quad E[\boldsymbol{w}(k)\boldsymbol{v}^T(l)] = \mathbf{0} \tag{11}$$

Assuming that the mathematical process model (8), the measured outputs $\boldsymbol{Y}(k) = \{\boldsymbol{y}(1)\boldsymbol{y}(2), ..., \boldsymbol{y}(k)\}$ (9), and the statistics (10), (11) are known, the filtering problem can then be formulated as determining the estimates

$$\begin{aligned} \hat{\boldsymbol{x}}(k|k-1) &= E[\boldsymbol{x}(k)|\boldsymbol{Y}(k-1)] \\ \hat{\boldsymbol{x}}(k|k) &= E[\boldsymbol{x}(k)|\boldsymbol{Y}(k)] \end{aligned} \tag{12}$$

and the associated error covariance matrices

$$\begin{aligned} \boldsymbol{P}(k|k-1) &= E[\,\delta\boldsymbol{x}(k|k-1)\,\delta\boldsymbol{x}^T(k|k-1)] \\ \boldsymbol{P}(k|k) &= E[\,\delta\boldsymbol{x}(k|k)\,\delta\boldsymbol{x}^T(k|k)] \end{aligned} \tag{13}$$

where $\delta\boldsymbol{x}(k|k-1) = \boldsymbol{x}(k) - \hat{\boldsymbol{x}}(k|k-1)$ is the one–step–ahead prediction error and $\delta\boldsymbol{x}(k|k) = \boldsymbol{x}(k) - \hat{\boldsymbol{x}}(k|k)$ is the filtering error. From this point on, the notations $\hat{\boldsymbol{x}}(i|j)$ and $\boldsymbol{P}(i|j)$ denote the state estimate and the associated covariance matrix given the available data $\boldsymbol{Y}(j)$, respectively.

The solution of the formulated problem (12)–(13) gives the optimal state estimator for the system defined by equations (8)–(11) and is described by the following set of equations (Anderson and Moore, 1979)

$$\hat{\boldsymbol{x}}(k+1|k) = \boldsymbol{A}(k)\hat{\boldsymbol{x}}(k|k) + \boldsymbol{B}(k)\boldsymbol{u}(k) \tag{14}$$

$$\boldsymbol{P}(k+1|k) = \boldsymbol{A}(k)\boldsymbol{P}(k|k)\boldsymbol{A}^T(k) + \boldsymbol{C}(k)\boldsymbol{Q}(k)\boldsymbol{C}^T(k) \tag{15}$$

$$\boldsymbol{V}(k+1) = \boldsymbol{H}(k+1)\boldsymbol{P}(k+1|k)\boldsymbol{H}^T(k+1) + \boldsymbol{R}(k+1) \tag{16}$$

$$\boldsymbol{K}(k+1) = \boldsymbol{P}(k+1|k)\boldsymbol{H}^T(k+1)\boldsymbol{V}^{-1}(k+1) \tag{17}$$

$$\boldsymbol{\nu}(k+1) = \boldsymbol{y}(k+1) - \boldsymbol{H}(k)\hat{\boldsymbol{x}}(k+1|k) \tag{18}$$

$$\hat{\boldsymbol{x}}(k+1|k+1) = \hat{\boldsymbol{x}}(k+1|k) + \boldsymbol{K}(k+1)\boldsymbol{\nu}(k+1) \tag{19}$$

$$\boldsymbol{P}(k+1|k+1) = \boldsymbol{P}(k+1|k) - \boldsymbol{K}(k+1)\boldsymbol{H}(k+1)\boldsymbol{P}(k+1|k) \tag{20}$$

The initial conditions $\hat{\boldsymbol{x}}(0|0)$ and $\boldsymbol{P}(0|0)$ are defined by (10). In the above equations, $\boldsymbol{\nu}(k+1)$ denotes the innovation sequence, which is an independent Gaussian random sequence with zero mean and known covariance matrix $\boldsymbol{V}(k+1)$ given by (16). This property of $\boldsymbol{\nu}(k+1)$ is very attractive for the FDI schemes. For stationary processes, i.e., when the matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{H}$, $\boldsymbol{Q}$, and $\boldsymbol{R}$ are constants, the filter covariance matrices ($\boldsymbol{P}(k+1|k)$ and $\boldsymbol{P}(k+1|k+1)$) approach a steady–state constant value $\boldsymbol{M}_s$ and $\boldsymbol{P}_s$, respectively. In these cases, only a constant filter gain $\boldsymbol{K}_s$ needs to be stored for on–line applications.

Unfortunately, the Kalman filter (14)–(20) can be designed correctly only if the mathematical model (8)–(11) describes the actual plant exactly without any system uncertainties. However, most real systems contain system uncertainties (variation of parameters, nonlinearities of the system, and "coloured" noises). Filters designed neglecting these system properties provide state estimates that can be confusing as to whether the observed behaviour is caused by the system uncertainties occurring in normal operation or by the faulty instrument. Hence, as it was pointed out earlier by many authors (Clark, 1975; Watanabe and Himmelblau, 1982; Frank, 1990), for such systems with uncertainties, robust filters should be designed, i.e., filters which are minimally sensitive to system uncertainties. The adaptive and nonlinear estimators are more complex than the linear one; however, they decrease the effect of modeling errors. It should be noticed that the sensitivity to modeling errors is the key problem in the application of FDI schemes based on analytical redundancy (Frank, 1990). Keeping this in mind, in the next part of our paper adaptive and robust algorithms of the Kalman filter will be discussed.

3. Adaptive Estimation

Assume that in our mathematical model (8)–(9) describing the actual plant, some of the elements in matrices $\boldsymbol{A}(k)$, $\boldsymbol{B}(k)$, and/or $\boldsymbol{H}(k)$ are unknown. Denoting these unknown elements by the vector $\boldsymbol{\theta}(k)$ ($dim\boldsymbol{\theta}(k) = s$), the mentioned matrices can be redenoted as $\boldsymbol{A}(\boldsymbol{\theta},k)$, $\boldsymbol{B}(\boldsymbol{\theta},k)$, and $\boldsymbol{H}(\boldsymbol{\theta},k)$. In general, the filtering problem for the linear system (8)–(11), but with unknown parameter vector $\boldsymbol{\theta}$, can be solved

by three different approaches (Anderson and Moore, 1979):
- parallel bank of Kalman filters (parallel processing),
- two–level estimation strategy (extended least squares),
- extended Kalman filtering (nonlinear estimation).

Among these approaches, the two–level estimation strategy is rarely applied in the FDI schemes (see Watanabe and Himmelblau, 1983). Thus, here we consider only the two other approaches.

3.1. Parallel Bank of Kalman Filters

The idea of parallel processing is very attractive in the solution of many problems both in fault detection (Frank, 1990) and filtering theory (Patton *et al.*, 1989). The application of the parallel approach to the adaptive estimation problem is described below.

Assume that the unknown parameter vector $\boldsymbol{\theta}$ is constant and can be discretized or suitably quantized to a finite number of grid points $\{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, ..., \boldsymbol{\theta}_M\}$ with assumed *a priori* probability for each $\boldsymbol{\theta}_i$, $i = 1, 2, ..., M$. Under such assumptions, the conditional mean state estimate $\hat{\boldsymbol{x}}(k+1|k)$ can be expressed using the conditional estimates $\hat{\boldsymbol{x}}(k+1|k;\ \boldsymbol{\theta}_i)$ and the conditional probabilities $p(\boldsymbol{\theta}_i|\boldsymbol{Y}(k+1))$ as

$$\hat{\boldsymbol{x}}(k+1|k) = \sum_{i=1}^{M} \hat{\boldsymbol{x}}(k+1|k;\ \boldsymbol{\theta}_i)\ p(\boldsymbol{\theta}_i|\boldsymbol{Y}(k+1)) \tag{21}$$

The conditional estimates $\hat{\boldsymbol{x}}(k+1|k;\ \boldsymbol{\theta}_i)$ are defined using conditional Kalman filters (14)–(20) for each of the parameter values $\boldsymbol{\theta}_i$, $i = 1, 2, ..., M$. The conditional probabilities $p(\boldsymbol{\theta}_i|\boldsymbol{Y}(k+1)$ are calculated recursively according to the following expression (Anderson and Moore, 1979)

$$p(\boldsymbol{\theta}_i|\boldsymbol{Y}(k+1)) = \tag{22}$$
$$c|\boldsymbol{V}^{-1}(k+1;\boldsymbol{\theta}_i)|^{\frac{1}{2}} \exp\left\{-\frac{1}{2}\boldsymbol{\nu}^T(k+1;\boldsymbol{\theta}_i)\boldsymbol{V}^{-1}(k+1;\boldsymbol{\theta}_i)\boldsymbol{\nu}(k+1;\boldsymbol{\theta}_i)\right\}p(\boldsymbol{\theta}_i;\boldsymbol{Y}(k))$$

where c is a normalizing constant independent of $\boldsymbol{\theta}_i$, chosen to ensure that $\sum_{i=1}^{M} p(\boldsymbol{\theta}_i|\boldsymbol{Y}(k+1)) = 1$, and $\boldsymbol{\nu}(k+1;\ \boldsymbol{\theta}_i)$ denotes the innovation sequence of the Kalman filter with the covariance $\boldsymbol{V}(k+1;\ \boldsymbol{\theta}_i)$ for $\boldsymbol{\theta}_i$. Thus, the adaptive estimator realizing this algorithm (21)–(22) consists of a bank of M parallel Kalman filters, each tuned to $\boldsymbol{\theta}_i(i = 1, 2, ..., M)$.

It is worth to note that this approach is very useful for fault detection and has been applied by Watanabe (see book: Patton *et al.*, 1989) for detecting sensor faults and estimating the state under faulty modes. The relatively high computing complexity of this algorithm can be considered a disadvantage; however, applying the decomposition approaches and the parallel computing methods, the computational cost can be considerably reduced.

3.2. Extended Kalman Filtering

In many applications of the fault detection and diagnosis, both the states and parameters are estimated via extended Kalman filtering (Himmelblau, 1978; Yoshimura *et al.*, 1979; Halme and Seikainaho, 1986; Chao and Paolella, 1990). In this technique, the unknown parameter vector $\boldsymbol{\theta}$ is augmented to the process state vector. In other words, our process model (8) and (9) with unknown parameter vector $\boldsymbol{\theta}$ can be described as follows

$$\boldsymbol{x}(k+1) = \boldsymbol{A}(\boldsymbol{\theta},k)\boldsymbol{x}(k) + \boldsymbol{B}(\boldsymbol{\theta},k)\boldsymbol{u}(k) + \boldsymbol{C}(k)\boldsymbol{w}(k) \tag{23}$$

$$\boldsymbol{\theta}(k+1) = \boldsymbol{\theta}(k) + \boldsymbol{w}_\theta(k) \tag{24}$$

$$\boldsymbol{y}(k) = \boldsymbol{H}(\boldsymbol{\theta},k)\boldsymbol{x}(k) + \boldsymbol{v}(k) \tag{25}$$

Equation (24) models the parameter dynamics by using a zero mean $s-$dimensional white Gaussian disturbance vector $\boldsymbol{w}_\theta(k)$ with the covarince given by $(s\times s)-$dimensional matrix $\boldsymbol{Q}_\theta(k)$. On defining the following augmented state vector $z(k)$

$$\boldsymbol{z}^T(k) \triangleq \left[\boldsymbol{x}^T(k), \boldsymbol{\theta}^T(k)\right] \tag{26}$$

equations (23)–(25) may be rewritten as

$$\boldsymbol{z}(k+1) = \boldsymbol{f}(\boldsymbol{z}(k),\boldsymbol{u}(k),k) + \boldsymbol{C}_z(k)\boldsymbol{w}_z(k) \tag{27}$$

$$\boldsymbol{y}(k) = \boldsymbol{h}(\boldsymbol{z}(k),k) + \boldsymbol{v}(k) \tag{28}$$

where

$$\boldsymbol{w}_z^T(k) \triangleq \left[w^T(k), w_\theta^T(k)\right],$$

and

$$\boldsymbol{C}_z(k) \triangleq \begin{bmatrix} \boldsymbol{C}(k) & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix}, \quad \boldsymbol{Q}_z(k) \triangleq \begin{bmatrix} \boldsymbol{Q}(k) & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Q}_\theta(k) \end{bmatrix}$$

$\boldsymbol{f}(\cdot)$ and $\boldsymbol{h}(\cdot)$ are $(n+s)-$ and $m-$dimensional known nonlinear function vectors, respectively.

Since the augmented model (27)–(28) is nonlinear, the filtering problem for such systems can be formulated as a direct nonlinear optimization problem (Seinfeld and Gavales, 1970; Kalogerakis and Luus, 1980), or can be solved by using the extended Kalman filter algorithm (Anderson and Moore, 1979). In summary, the solution procedure of the first approach requires long computation times while the latter is more attractive in applications (Himmelblau, 1978; Ishii *et al.*, 1980; Yoshimura *et al.*, 1980), and will be presented in the next section.

4. Nonlinear Filtering

Consider the nonlinear state estimation problem for the discrete model described by equations (27)–(28) with known statistics (10)–(11). Different nonlinear techniques for solving this problem are available (Anderson and Moore, 1979; Sorenson, 1985) and a short survey of the recursive state estimation techniques is given by Misawa and Hendrick (1988). Among these techniques, the extended Kalman filter method is widely used by most investigators to solve practical problems (Sorenson, 1985). Therefore, this suboptimal filter for the process (27)–(28) will be presented below.

As a result of applying different approaches to derive the extended Kalman filter (Jazwinski, 1970; Anderson and Moore, 1979; Sorenson, 1985), the following system of equations can be obtained

$$\hat{\boldsymbol{z}}(k+1|k) = \boldsymbol{f}(\hat{\boldsymbol{z}}(k|k), \boldsymbol{u}(k), k) \tag{29}$$

$$\boldsymbol{P}(k+1|k) = \boldsymbol{A}_f(k)\boldsymbol{P}(k|k)\boldsymbol{A}_f^T(k) + \overline{\boldsymbol{Q}}_z(k) \tag{30}$$

$$\boldsymbol{V}(k+1) = \boldsymbol{H}_h(k+1)\boldsymbol{P}(k+1|k)\boldsymbol{H}_h^T(k+1) + \boldsymbol{R}(k+1) \tag{31}$$

$$\boldsymbol{K}(k+1) = \boldsymbol{P}(k+1|k)\boldsymbol{H}_h^T(k+1)\boldsymbol{V}^{-1}(k+1) \tag{32}$$

$$\boldsymbol{\nu}(k+1) = \boldsymbol{y}(k+1) - \boldsymbol{h}(\hat{\boldsymbol{z}}(k+1|k), k+1) \tag{33}$$

$$\hat{\boldsymbol{z}}(k+1|k+1) = \hat{\boldsymbol{z}}(k+1|k) + \boldsymbol{K}(k+1)\boldsymbol{\nu}(k+1) \tag{34}$$

$$\boldsymbol{P}(k+1|k+1) = \boldsymbol{P}(k+1|k) - \boldsymbol{K}(k+1)\boldsymbol{H}_h(k+1)\boldsymbol{P}(k+1|k) \tag{35}$$

where the initialization is provided by

$$\hat{\boldsymbol{z}}(0|0) = \hat{\boldsymbol{z}}_0, \quad \boldsymbol{P}(0|0) = \boldsymbol{P}_0$$

In equations (29)–(35), the following notations have been used

$$\boldsymbol{A}_f = \left.\frac{\partial \boldsymbol{f}(\boldsymbol{z}(k), \boldsymbol{u}(k), k)}{\partial \boldsymbol{z}}\right|_{\boldsymbol{z}=\hat{\boldsymbol{z}}(k|k)}, \quad \boldsymbol{H}_h(k) = \left.\frac{\partial \boldsymbol{h}(\boldsymbol{z}(k), k)}{\partial \boldsymbol{z}}\right|_{\boldsymbol{z}=\hat{\boldsymbol{z}}(k+1|k)} \tag{36}$$

and

$$\overline{Q}_z(k) = C_z(k)Q_z(k)C_z^T(k)$$

One should notice that the above extended Kalman filter (29)–(35) has the same structure as a linear one (14)–(20), but now the notations $\hat{z}(k+1|k)$ and $P(k+1|k)$ denote approximate conditional means and covariances, respectively. In this case, equations (29)–(36) are coupled because matrices $A_f(k)$ and $H_h(k)$ are functions of $\hat{z}(k+1|k)$ and should be computed on-line. The lack of guaranteed robustness and difficulties in implementation are considered as the main drawbacks of such an algorithm, even though it is widely used (Sorensen, 1985). Therefore, to overcome some of these drawbacks, different improvements have been proposed. For example, as suggested by Safonov and Athans (1978) and implemented by Chao and Paolella (1990), the constant gain extended Kalman filter allows us to reduce the substantial real–time computational burden imposed by the algorithm (29)–(35). Another attractive approach that allows one to alleviate the computational difficulties and provide robustness in the case of uncertainty in the mathematical model (27)–(28) has been proposed by Krasovsky (1976) and then extended to the distributed parameter systems by Korbicz (1986).

5. Suboptimal Nonlinear Filtering

The main idea of the suboptimal approach can be formulated as follows. By definition, the one–step ahead prediction error covariance matrix is defined by

$$P(k+1|k) = E[\,\delta z(k+1|k)\,\delta z^T(k+1|k)] \tag{37}$$

where $\delta z(k+1|k) = z(k+1) - \hat{z}(k+1|k)$ is the one–step ahead prediction error. As we know, this matrix is described by equation (15) in the linear case, and by equation (30) on applying the extended Kalman filter for the nonlinear estimation problem. In many cases, the matrix $P(k+1|k)$ can be approximated by using the following formula

$$\widehat{P}(k+1|k) = \frac{1}{k+1}\sum_{l=1}^{k+1} \delta z(l+1|l)\,\delta z^T(l+1|l) \tag{38}$$

where $\widehat{P}(k+1|k)$ denotes the estimate of $P(k+1|k)$.

As a result of simple transformation of (38), the estimate $\widehat{P}(k+1|k)$ can be defined by the following recursive expression (Korbicz, 1986)

$$\widehat{P}(k+1|k) = \widehat{P}(k|k) + \gamma(k)\Big\{\,\delta z(k+1|k)\,\delta z^T(k+1|k) - \widehat{P}(k|k)\Big\} \tag{39}$$

where the sequence $\gamma(k)$ is chosen from some convergence conditions.

Taking into account the approximate solution (39), the suboptimal filtering algorithm for the nonlinear system (27)–(28) is implemented as (Korbicz and Zgurovsky, 1991)

$$\hat{z}(k+1|k) = f(\hat{z}(k|k), u(k), k) \quad (40)$$

$$\delta z(k+1|k) = A_f(k)\, \delta z(k|k) + C_z(k) w_z(k) \quad (41)$$

$$\hat{P}(k+1|k) = \hat{P}(k|k) + \gamma(k)\left[\delta z(k+1|k)\delta z^T(k+1|k) - \hat{P}(k|k)\right] \quad (42)$$

$$V(k+1) = H_h(k+1)\hat{P}(k+1|k)H_h^T(k+1) + R(k+1) \quad (43)$$

$$K(k+1) = \hat{P}(k+1|k)H_h^T(k+1)V^{-1}(k+1) \quad (44)$$

$$\nu(k+1) = y(k+1) - h(\hat{z}(k+1|k), k+1) \quad (45)$$

$$\hat{z}(k+1|k+1) = \hat{z}(k+1|k) + K(k+1)\nu(k+1) \quad (46)$$

$$\delta z(k+1|k+1) = \delta z(k+1|k) - K(k+1)\nu(k+1) \quad (47)$$

$$\hat{P}(k+1|k+1) = \hat{P}(k+1|k) - K(k+1)H_h(k+1)\hat{P}(k+1|k) \quad (48)$$

with initial conditions, $\hat{z}(0|0) = \hat{z}_0$, $\hat{P}(0|0) = P_0$.

In comparison with the extended Kalman filter (29)–(35), the suboptimal algorithm (40)–(48), is described by the additional equations for the prediction error (41) and the filtering error (47). As the covariance matrix estimate $\hat{P}(k+1|k)$ is defined with respect to measurement data, this suboptimal estimator (40)–(48) is less sensitive to the incompleteness of *a priori* data about the actual system. It is a robust algorithm and therefore can be implemented both for nonlinear and linear discrete–time systems.

6. Kalman Filter for Nonlinear Systems with Coupled Static and Dynamic Models

In practice, actual systems are often described by coupled static and dynamic models. Thus it is important to consider the application of the extended Kalman filter algorithm to the joint parameter and state estimation for systems with coupled static and dynamic models.

Let the model of a general stochastic system with unknown parameter vector θ be described mathematically by the following equations (Fathi *et al.*, 1991)

$$x_d(k+1) = f_d(x_d(k), x_s(k), \theta(k), u(k), k) + w_d(k) \quad (49)$$

$$\begin{aligned} \mathbf{0} &= \boldsymbol{f}_s(\boldsymbol{x}_d(k+1), \boldsymbol{x}_s(k+1), \boldsymbol{\theta}(k+1), \boldsymbol{u}(k+1), k+1) + \boldsymbol{w}_s(k+1) &(50)\\ \boldsymbol{y}(k+1) &= \boldsymbol{h}(\boldsymbol{x}_d(k+1), \boldsymbol{x}_s(k+1), \boldsymbol{\theta}(k+1), k+1) + \boldsymbol{v}(k+1) &(51)\end{aligned}$$

where equation (49) describes the dynamic model, and equation (50) describes the steady–state (static) model. In (49)–(51) sub–indices d and s denote variables and parameters associated with slow and fast dynamics, respectively. All remaining notations are the same as in the previous sections.

Assuming that the true parameter vector $\boldsymbol{\theta}$ varies according to

$$\boldsymbol{\theta}(k+1) = \boldsymbol{\theta}(k) + \boldsymbol{w}_\theta(k) \tag{52}$$

the augmented state vector $\boldsymbol{z}(k)$, and the noise vector $\boldsymbol{w}(k)$ can be defined as

$$\begin{aligned} \boldsymbol{z}^T(k) &\triangleq \left[\boldsymbol{x}_d^T(k), \boldsymbol{x}_s^T(k), \boldsymbol{\theta}^T(k)\right] \\ \boldsymbol{w}^T(k) &\triangleq \left[\boldsymbol{w}_d^T(k), \boldsymbol{w}_s^T, \boldsymbol{w}_\theta^T(k)\right] \end{aligned} \tag{53}$$

Furthermore, we assume that we have the statistical information for all the random variables, which are Gaussian sequences and uncorrelated.

Then, the modified extended Kalman filter for this problem can be defined by the following system of equations (Fathi *et al.*, 1991)

$$\begin{aligned}
\widehat{\boldsymbol{x}}_d(k+1|k) &= \boldsymbol{f}_d(\widehat{\boldsymbol{x}}_d(k|k), \widehat{\boldsymbol{x}}_s(k|k), \widehat{\boldsymbol{\theta}}(k|k), \boldsymbol{u}(k), k) &(54)\\
\widehat{\boldsymbol{x}}_s(k+1|k) &= \mathrm{Sol}_{\widehat{\boldsymbol{x}}_s}\{\boldsymbol{f}_s(\widehat{\boldsymbol{x}}_d(k+1|k), \widehat{\boldsymbol{x}}_s(k+1|k), \\
&\qquad \widehat{\boldsymbol{\theta}}(k+1|k), \boldsymbol{u}(k+1), k+1) = \mathbf{0}\} &(55)\\
\widehat{\boldsymbol{\theta}}(k+1|k) &= \widehat{\boldsymbol{\theta}}(k|k) &(56)\\
\boldsymbol{\nu}(k+1) &= \boldsymbol{y}(k+1) - \boldsymbol{h}(\widehat{\boldsymbol{z}}(k+1|k), k+1) &(57)\\
\boldsymbol{P}(k+1|k) &= \boldsymbol{A}_f(k)\boldsymbol{P}(k|k)\boldsymbol{A}_f^T(k) + \boldsymbol{Q}_f(k) &(58)\\
\boldsymbol{V}(k+1) &= \boldsymbol{H}_h(k+1)\boldsymbol{P}(k+1|k)\boldsymbol{H}_h^T(k+1) + \boldsymbol{R}(k+1) &(59)\\
\boldsymbol{K}(k+1) &= \boldsymbol{P}(k+1|k+1)\boldsymbol{H}_h^T(k+1)\boldsymbol{R}^{-1}(k+1) &(60)\\
\widehat{\boldsymbol{z}}(k+1|k+1) &= \widehat{\boldsymbol{z}}(k+1|k) + \boldsymbol{K}(k+1)\boldsymbol{\nu}(k+1) &(61)\\
\boldsymbol{P}(k+1|k+1) &= \boldsymbol{P}(k+1|k) - \boldsymbol{P}(k+1|k)\boldsymbol{H}_h^T(k+1)\boldsymbol{V}^{-1}(k+1) \times \\
&\qquad \boldsymbol{H}_h(k+1)\boldsymbol{P}(k+1|k) &(62)
\end{aligned}$$

with initial conditions, $\widehat{\boldsymbol{x}}(0|0) = \widehat{\boldsymbol{x}}_0$ and $\boldsymbol{P}(0|0) = \boldsymbol{P}_0$. $\mathrm{Sol}_{\widehat{\boldsymbol{x}}_s}$ denotes the solution of the static equations for the vector $\widehat{\boldsymbol{x}}_s$. It should be noticed that $\widehat{\boldsymbol{z}}^T(k+1|k) \triangleq \left[\widehat{\boldsymbol{x}}_d^T(k+1|k), \widehat{\boldsymbol{x}}_s^T(k+1|k), \boldsymbol{\theta}^T(k+1|k)\right]$.

The submatrices $\boldsymbol{A}_{ij}$ of the matix $\boldsymbol{A}_f(k) = \left\{ \vdots \boldsymbol{A}_{ij} \vdots \right\}$ for $i,j = 1,2,3$ are given by

$$\boldsymbol{A}_{11} = \frac{\partial \widehat{\boldsymbol{f}}_d}{\partial \boldsymbol{x}_d} \qquad \boldsymbol{A}_{12} = \frac{\partial \widehat{\boldsymbol{f}}_d}{\partial \boldsymbol{x}_s} \qquad \boldsymbol{A}_{13} = \frac{\partial \widehat{\boldsymbol{f}}_d}{\partial \boldsymbol{\theta}}$$

$$\boldsymbol{A}_{21} = -\left(\frac{\partial \overline{\boldsymbol{f}}_s}{\partial \boldsymbol{x}_s}\right)^{-1} \frac{\partial \overline{\boldsymbol{f}}_s}{\partial \boldsymbol{x}_d} \frac{\partial \widehat{\boldsymbol{f}}_d}{\partial \boldsymbol{x}_d} \qquad \boldsymbol{A}_{22} = -\left(\frac{\partial \overline{\boldsymbol{f}}_s}{\partial \boldsymbol{x}_s}\right)^{-1} \frac{\partial \overline{\boldsymbol{f}}_s}{\partial \boldsymbol{x}_d} \frac{\partial \widehat{\boldsymbol{f}}_d}{\partial \boldsymbol{x}_s} \tag{63}$$

$$\boldsymbol{A}_{23} = -\left(\frac{\partial \overline{\boldsymbol{f}}_s}{\partial \boldsymbol{x}_s}\right)^{-1} \left(\frac{\partial \overline{\boldsymbol{f}}_s}{\partial \boldsymbol{x}_d} \frac{\partial \widehat{\boldsymbol{f}}_d}{\partial \boldsymbol{\theta}} + \frac{\partial \overline{\boldsymbol{f}}_s}{\partial \boldsymbol{\theta}}\right)$$

$$\boldsymbol{A}_{31} = \boldsymbol{0} \qquad \boldsymbol{A}_{32} = \boldsymbol{0} \qquad \boldsymbol{A}_{33} = \boldsymbol{I}$$

The submatrices $\boldsymbol{Q}_{ij}$ of the matrix $\boldsymbol{Q}_f(k) = \left\{ \vdots \boldsymbol{Q}_{ij} \vdots \right\}$ for $i,j = 1,2,3$ are defined by the following expressions

$$\boldsymbol{Q}_{11} = \boldsymbol{Q}_d(k) \qquad \boldsymbol{Q}_{12} = -\boldsymbol{Q}_d(k) \left(\frac{\partial \overline{\boldsymbol{f}}_s}{\partial \boldsymbol{x}_d}\right)^T \left(\frac{\partial \overline{\boldsymbol{f}}_s}{\partial \boldsymbol{x}_s}\right)^{-T} \qquad \boldsymbol{Q}_{13} = \boldsymbol{0}$$

$$\boldsymbol{Q}_{21} = -\left(\frac{\partial \overline{\boldsymbol{f}}_s}{\partial \boldsymbol{x}_s}\right)^{-1} \frac{\partial \overline{\boldsymbol{f}}_s}{\partial \boldsymbol{x}_d} \boldsymbol{Q}_d(k) \qquad \boldsymbol{Q}_{23} = -\left(\frac{\partial \overline{\boldsymbol{f}}_s}{\partial \boldsymbol{x}_s}\right)^{-1} \frac{\partial \overline{\boldsymbol{f}}_s}{\partial \boldsymbol{\theta}} \boldsymbol{Q}_\theta(k)$$

$$\boldsymbol{Q}_{22} = \left(\frac{\partial \overline{\boldsymbol{f}}_s}{\partial \boldsymbol{x}_s}\right)^{-1} \left[\boldsymbol{Q}_s(k) + \left(\frac{\partial \overline{\boldsymbol{f}}_s}{\partial \boldsymbol{x}_d}\right) \boldsymbol{Q}_d(k) \left(\frac{\partial \overline{\boldsymbol{f}}_s}{\partial \boldsymbol{x}_d}\right)^T + \left(\frac{\partial \overline{\boldsymbol{f}}_s}{\partial \boldsymbol{\theta}}\right) \boldsymbol{Q}_\theta(k) \left(\frac{\partial \overline{\boldsymbol{f}}_s}{\partial \boldsymbol{\theta}}\right)^T\right] \left(\frac{\partial \overline{\boldsymbol{f}}_s}{\partial \boldsymbol{x}_s}\right)^{-T}$$

$$\boldsymbol{Q}_{31} = \boldsymbol{0} \qquad \boldsymbol{Q}_{32} = -\boldsymbol{Q}_\theta(k) \left(\frac{\partial \overline{\boldsymbol{f}}_s}{\partial \boldsymbol{\theta}}\right)^T \left(\frac{\partial \overline{\boldsymbol{f}}_s}{\partial \boldsymbol{x}_s}\right)^{-T} \qquad \boldsymbol{Q}_{33} = \boldsymbol{Q}_\theta(k) \tag{64}$$

where

$$\frac{\partial \overline{\boldsymbol{f}}_s}{\partial \boldsymbol{z}} = \left.\frac{\partial \boldsymbol{f}_s}{\partial \boldsymbol{z}}\right|_{\boldsymbol{z}=\widehat{\boldsymbol{z}}(k+1|k)} \qquad \frac{\partial \widehat{\boldsymbol{f}}_d}{\partial \boldsymbol{z}} = \left.\frac{\partial \boldsymbol{f}_d}{\partial \boldsymbol{z}}\right|_{\boldsymbol{z}=\widehat{\boldsymbol{z}}(k+1|k+1)}$$

and $(\cdot)^{-T}$ denotes the transposed inverse of $(\cdot)$.

The effectiveness of such a modified extended Kalman filter algorithm was shown by Fathi *et al.* (1991) for the solution of the estimation problem for a complex nuclear reactor system. Nevertheless, the extended Kalman filter used to solve the joint state and parameter estimation for the dynamic or the coupled static and dynamic systems has some drawbacks, which should be attenuated through additional improvements for applications in fault detection and isolation systems.

7. Correction of the Extended Kalman Filter

From many applications (Sorenson, 1985), it is known that the extended Kalman filter can respond properly to parameter variations if they are relatively slow with time. However, in fault detection systems, the variations of unknown parameters can be abrupt and in addition can occur any time. In that case, the estimation of unknown parameters becomes difficult since the covariance matrix of estimation errors in the extended Kalman filter algorithms for the unknown parameters decreases monotonically, and thus the filter cannot effectively estimate the parameter changes occuring later in time. To prevent such filter degradation, several techniques have been proposed by Jazwinski (1970) and Yoshimura *et al.*(1979) in which the monotonical decrease of the filter gain is prevented by additional conditions.

In general, this filter degradation problem is solved in the following way. First, a new condition is checked and then the extended Kalman filter is modified. For instance, Yoshimura *et al.*(1979) proposes to check the following condition

$$\left|\theta_i^n - \widehat{\theta}_i(k|k)\right| > d(P_{\theta i}(k+1|k))^{1/2}, \quad i = 1, 2, ..., s \tag{65}$$

where θ_i^n and $\widehat{\theta}_i(k|k)$ are nominal values and estimates of θ_i, respectively. $P_{\theta i}(k+1|k)$ are diagonal elements of the error variances $\boldsymbol{P}_\theta(k+1|k)$ and d denotes a positive constant. If the condition (65) is satisfied for one or more parameter estimates, then the modified variances $P_{\theta i}^m(k+1|k)$ are redefined as

$$P_{\theta i}^m(k+1|k) = \left[\theta_i^n - \widehat{\theta}_i(k|k)\right]^2 \Big/ d^2 \tag{66}$$

and are substituted into the filter equation. In addition, after such modification, new values of θ_i^n are changed as $\theta_i^n = \widehat{\theta}_i(k|k)$.

Another filter gain modification was proposed by Sriyananda (1972) and then extended to the discrete–time distributed parameter systems (Korbicz, 1985). To detect divergence in the Kalman filter algorithm, the following condition is tested

$$\boldsymbol{\nu}^T(k+1)\,\boldsymbol{\nu}(k+1) \leq \gamma \mathrm{Trace}\{\boldsymbol{V}(k+1)\} \tag{67}$$

where the left side of this inequality is defined by using the innovation sequence from the filter while the theoretical covariance matrix of the innovation sequence $\boldsymbol{V}(k+1)$ is used on the right side of (67). In equation (67) γ denotes a positive constant ($\gamma \leq 1.2$). If during the execution of the Kalman filter algorithm the condition (67) is satisfied, then the covariance matrix $\boldsymbol{P}(k+1|k)$ is modified as follows

$$\boldsymbol{P}^m(k+1|k) = s(k+1)\boldsymbol{P}(k+1|k) \tag{68}$$

where the scalar correction coefficient $s(k+1)$ can be a positive constant ($s > 1$) or can be defined as a more complex function (Kuzovkov *et al.*, 1978; Korbicz, 1985).

Here it should be pointed out that the suboptimal filtering algorithm presented in section 5 does not have any modifications for the filter gain or equivalently for the prediction covariance matrix. However, on the contrary to the extended Kalman filter algorithm, in the suboptimal case the covariance matrix is defined with respect to the real estimation error, or in other words, with respect to the measurement data.

8. Conclusions

In this paper, a brief review of linear and nonlinear state estimation methods has been presented with respect to their application for fault detection and diagnosis using analytical redundancy and knowledge–based techniques. It is shown that there exist various state and parameter estimation methods (such as linear Kalman filter, extended Kalman filter, and adaptive filtering approach) that are often used in designing FDI schemes. These methods are basic in state estimation theory. However, the designed FDI systems can be more efficient by using the nonlinear filtering algorithms for dynamic systems or for systems with coupled static and dynamic models, or the suboptimal robust filtering algorithms for linear and nonlinear systems.

It should be noted that this paper has focused on state estimation methods used in FDI schemes, and implementational problems of these methods have not been discussed. We refer the reader to the overview papers

(Willsky, 1976;Frank, 1990) or to the book by Patton *et al.* (1989), for a detailed description, where the various known techniques for FDI systems design have been classified and discussed. For example, Willsky (1976) divided these techniques into the following groups: failure–sensitive filters, voting systems, jump process formulation, and innovation–based detection systems. The key problem in designing FDI systems which should be considered in all cases is the robustness problem. Therefore, taking into account the rapid development of the knowledge engineering for industrial expert systems (Johannsen and Alty, 1991), it seems that the robustness problem can be effectively solved by using coupled analytical and knowledge–based techniques. Recently, many simulation studies and experimental results have shown that FDI systems using simultaneously the analytical and knowledge–based redundancy are more powerful and more flexible (see Tzafestas, 1989). Unfortunately, at present most designed FDI systems do not use the power of analytical redundancy techniques. In our opinion, the estimation system conducts important and crucial information, which is of substantial assistance, to the expert system.

Furthermore, it should be noted that in the FDI schemes for large-scale systems, the computational burden is often a very crucial problem. In general, in the analytical model–based approach, this problem can be solved by using one of many different methods to decentralize estimation in large–scale systems (Gardner, 1989; Watanabe, 1989). Another approach that can reduce computational complexity is the use of an expert system accompanied by system decomposition. For each of the components of the decomposed system, individual estimators can be designed and then a set of rules in the knowledge base allows us to coordinate the information flow from each of the local estimators and make decisions about global estimators.

References

Anderson B.D.O. and Moore J.B. (1979): *Optimal Filtering.*–New York: Prentice Hall, Englewood Cliffs.

Basseville M. (1988): *Detecting changes in signals and systems. A survey.*– Automatica, v.24, No.3, pp.309–326.

Basseville M. and Benveniste A. (Eds) (1986): *Detection of Abrupt Changes in Signals and Dynamical Systems.*–Berlin: Springer–Verlag, (Ser. Lectures Notes on Control and Information Sciences, 1986, No.77).

Bryson J.A.E. and Ho Y.C. (1975): *Applied Optimal Control.*–New York: Hemisphere Publishers.

Chao D. and Paolella P. (1990): *Model-based failure detection and isolation of automotive powertrain systems* .–Proc. Amer. Control Conf., San Diego, May, v.3, pp.2888–2905.

Clark R.N., Fosth D.C. and Walton V.M. (1975): *Detection instrument malfunctions in control systems.*–IEEE Trans. Aerospace Electron. Syst., v.AES–11, pp.465–473.

Deckert J.C., Desai M.N., Deyst J.J. and Willsky A.S. (1977): *DFBW sensor failure identification using analytic redundancy.*–IEEE Trans. Automat. Control, v.AC–22, pp.795–809.

Davis M.H.A. (1975): *The application of nonlinear filtering to fault detection in linear systems.*–IEEE Trans. Automat. Control, v.AC–20, pp.257–259.

Eckert S., Loparo K. and Roth Z. (1986): *An application of nonlinear filtering to instrument failure detection in a pressurized water reactor.*–Nuclear Technology, v.74, No.2, pp. 139–151.

Fathi Z., Ramirez W.F. and Aarna O. (1991): *Joint state and parameter estimation for a nuclear reactor with fast and slow modes.*– Optimal Control Applications and Methods. (will appear).

Frank P.M. (1990): *Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy. A survey and some new results.*–Automatica, v.26, No.3, pp.459–474.

Gardner W.T. (1989): *Gain transfer: an algorithm for decentralized hierarchical estimation.* In *Control and Dynamic Systems. Advances in Theory and Application.*–Leondes C.T.(Ed), v.30, part 3 of 3, pp.19–64.–San Diego: Academic Press.

Gertler J.J., Fang X. and Luo Q. (1990): *Detection and diagnosis of plant failures: The orthogonal parity equation approach.* In *Control and Dynamic Systems. Advances in Theory and Applications.*– Leondes C.T.(Ed.), pp.159–216.–New York: Academic Press.

Gertler J. and Singer D. (1990): *A new structural framework for parity equation-based failure detection and isolation.* – Automatica, v.26, No.2, pp.381–388.

Haesloop D. and Holt B.R. (1990): *A neural network structure for system identification.*–Proc. Amer. Control Conf., San Diego, USA, May 23–25, v.3, pp.2460–2465.

Halme A. and Seikainaho J. (1986): *An adaptive filtering based method to detect sensor/actuator faults.*–Proc. IFAC Workshop, "Fault Detection and Safety in Chemical Plants," Kyoto, Japan, pp. 158–161.

Hengy D. and Frank P.M. (1986): *Component failure detection via nonlinear state observers.*–Proc. IFAC Workshop, "Fault Detection and Safety in Chemical Plants," Kyoto, Japan, pp.153–157.

Himmelblau D.M. (1978): *Fault Detection and Diagnosis in Chemical and Petrochemical Processes.*–New York: Elsevier.

Himmelblau D.M. (1986): *Fault detection and diagnosis– today and tomorrow.*–Proc. IFAC Workshop, "Fault Detection and Safety in Chemical Plants," Kyoto, Japan, pp.95–105.

Isermann R. (1984): *Process fault detection based on modeling and estimation methods. A survey.*–Automatica, v.20, No.4, pp.387–404.

Ishii N., Iwata A. and Suzumura N. (1980): *Detection of abrupt change in trend in the time series.*–Int. J. Syst. Sci., v.11, pp.557–566.

Jazwinski A.H. (1970): *Stochastic Processes and Filtering Theory.*–New York: Academic Press.

Johannsen G. and Alty J.L. (1991): *Knowledge engineering for industrial experts systems.*–Automatica, v.27, No.1, pp.97–114.

Jones H.L. (1973): *Failure Detection in Linear Systems.*–Ph.D. Thesis, MIT, Cambridge, MA.

Kalman R.E. (1960): *On a new approach to filtering and prediction problems.*–Trans. ASME J. Basic Engineering, v.82–D, No.1, pp.35–45.

Kalogerakis N. and Luus R. (1980): *Effective data length on the region of convergence in parameter estimation using quasilinearization.*–AIChE Journal, v.24, pp.670–672.

Kerr T.H. (1982): *False alarm and correct detection probabilities over a time interval for restricted classes of failure detection algorithms.*–IEEE Trans. Information Theory, v.IT–28, No.4, pp.619–631.

Kitamura M. (1980): *Detection of sensor failures in nuclear plant using analytic redundancy.*–Trans. Amer. Nucl. Sci., v.34, pp.581–583.

Korbicz J. (1985): *State estimation in discrete-time parameter systems under incomplete priori information about the system.*–Kybernetika, v.21, No.6, pp.470–479.

Korbicz J. (1986): *Suboptimal state estimation algorithm for the nonlinear discrete-time distributed parameter systems.*– Int. J. Syst. Sci., v.17, No.5, pp.725–734.

Korbicz J. and Zgurovsky M.Z. (1991): *Estimation and Control of Stochastic Distributed Parameter Systems.*–Polish Scientific Publishers, Warsaw (in Polish, in print).

Krasovsky A.A. (1976): *Adaptive algorithm of suboptimal estimation.*–Dokl. Acad. Sciences USSR, v.230, No.3, (in Russian).

Kuzovkov N.T., Karabanov S.V. and Salychev O.S. (1978): *Continuous and Discrete–Time Control Systems and Identification Methods.*- -Moscow: Mashinostrojenie.

Loparo K.A., Buchner M.R. and Vasudeva K.S. (1991): *Leak detection in an experimental heat exchanger process: A multiple model approach.*–IEEE Trans. Automat. Control, v.36 No.2, pp.167–177.

Loparo K.A., Roth Z.S. and Eckert S.J. (1986): *Nonlinear filtering for systems with random structure.*–IEEE Trans. Automat. Control, v.AC-31, No.1, pp.1064–1067.

Luck R. and Ray A. (1991): *Failure detection and isolation of ultrasonic ranging sensors for robotic applications* .–IEEE Trans. Syst. Man and Cybernetics, v.21, No.1, pp.221–227.

Mehra R.K. and Peschon J. (1971): *An innovations approach to fault detection and diagnosis in dynamical systems.* –Automatica, v.7, pp.637–640.

Miller T., Sutton R.S. and Werbos P.J. (1990): *Neural Networks for Control.*–Cambridge: MIT Press, MA.

Misawa E.A. and Hedrick J.K. (1988): *Nonlinear observers: A state–of–the–art Survey.* In *Recent Advances in Control of Nonlinear and Distributed Parameter Systems, Robust Control, and Aerospace Control Applications.*– Benstsman J. and Joshi S.M.(Eds.), New York, pp.59–69.

Naidu S.R., Zafiriou E. and McAvoy T.J. (1990): *Use of neural networks for sensor failure detection in a control system.*–IEEE Control Syst. Mag., v.10, No.3, pp.49–55.

Neumann D. (1990): *A knowledge based fault diagnosis system for the supervision of periodically and intermittent working machine tools.*–Proc. 11th IFAC Congress, Tallinn, USSR, pp.135–140.

Patton R., Frank P. and Clark R. (1989): *Fault Diagnosis in Dynamic Systems. Theory and Applications.*–New York: Prentice Hall.

Ramirez W.F. (1987): *Optimal state and parameter identification–an application to batch fermentation.*–Chem. Eng. Sci., v.42, pp.2749–2756.

Safonov M.G. and Athans M. (1978): *Robustness and computational aspects of nonlinear stochastic estimators and regulators.*- IEEE Trans. Automat. Control, v.AC-23, No.4, pp.717–725.

Seinfeld J.H. and Gavalas G.R. (1970): *Analysis of kinetic parameters from batch and integral reaction experiments.*–AIChE Journal, v.16, pp.644–647.

Sorenson H.W. (Ed.) (1985): *Kalman Filtering: Theory and Application.*–New York: IEEE Press.

Sriyananda H. (1972): *A simple method for the control of divergence in Kalman filter algorithms.*–Int. J. Control, v.16, No.6, pp.1101–1106.

Thompson H.A. and Fleming P.J. (1990): *Transputer–based fault tolerant strategies for a gas turbine engine controller.*–Proc. of Amer. Control Conf., San Diego, USA, May 23–25, v.3, pp.2918–2923.

Tzafestas S.G. (Ed.) (1989): *Knowledge–Based System Diagnosis, Supervision, and Control.*–New York–London: Plenum.

Venkatasubramanian V. and Chen K. (1989): *A neural network methodology for process fault diagnosis.*–AIChE Journal, v.35, pp.1993–2002.

Watanabe K. (1989): *Multiple model adaptive filtering approach to fault diagnosis in stochastic systems.* In *Fault Diagnosis in Dynamic Systems*, **Patton R., Frank P. and Clark R.** (Eds.).–New York: Prentice Hall, pp.411–438.

Watanabe K., Fukuda T. and Tzafestas S.G. (1991): *Learning algorithms of layered neural networks via extended Kalman filters.*–Int. J. Syst. Sci., v.22, No.4, pp.753–768.

Watanabe K. and Himmelblau D.M. (1982): *Instrument fault detection in systems with uncertainties.*–Int. J. Syst. Sci., v.13, No.2, pp.137–158.

Watanabe K. and Himmelblau D.M. (1983): *Fault diagnosis in nonlinear chemical processes.*–Part I. *Theory.*–AIChE Journal, v.29, No.2, pp.243–249. Part II. *Application to a chemical reactor.*–AIChE Journal, v.29, No.2, pp.250– 260.

Watanabe K., Matsuura I., Abe M. and Kubota M. (1989): *Incipient fault diagnosis of chemical processes via artificial neural networks.*–AIChE Journal, v.35, pp.1803–1812.

Willsky A.S. (1976): *A survey of design methods for failure detection in dynamic systems.*–Automatica, v.12, pp.601–611.

Willsky A.S., Deyst J.J. and Crawford B.S. (1974): *Adaptive filtering and self–test methods for failure detection and compensation.*–Proc. of the 19th JACC, Austin, Texas, USA, June 19–21, pp.637–649.

Wunnenberg J. and Frank P.M. (1990): *Dynamic model–based incipient fault detection concept for robot.*–Proc. 11th IFAC World Congress, Tallinn, Estonia, USSR, pp. 1420–1426.

Yao S.Ch. and Zafiriou E. (1990): *Control system sensor failure detection via networks of localized receptive fields.*–Proc. Amer. Control Conf., San Diego, USA, May 23–25, v.3, pp.2472–2477.

Yoshimura T., Konishi K., Kiyozumi R. and Soeda T. (1980): *Identification of unknown parameters in linear discrete-time systems by a modified extended Kalman filter.*–Int. J. Syst. Sci., v.11, pp.97–105.

Yoshimura T., Watanabe K., Konish K. and Soeda T. (1979): *A sequential failure detection approach and the identification of failure parameters.*–Int. J. Syst. Sci., v.10, No.7, pp.827–836.

This paper was written during Korbicz's stay at the Department of Chemical Engineering at the University of Colorado, U.S.A., sponsored by IREX.