1

Nonlinear Methods for Fault Diagnosis

Silvio SIMANI[1] and Paolo CASTALDI[2]
[1] *Department of Engineering, University of Ferrara, Emilia-Romagna, Italy*
[2] *Department of Electrical, Electronic and Information Engineering, University of Bologna, Italy*

1.1. Introduction

The model-based approach to fault diagnosis in technical processes has been receiving more and more attention over the last four decades, in the contexts of both research and real plant application.

Stemming from this activity, a large number of methods can be found in current literature based on the use of mathematical models of the technical process under diagnosis and on exploiting advanced control theory.

Model-based fault diagnosis methods usually use residuals that indicate changes between the process and the model. One general assumption is that the residuals are changed significantly so that detection is possible. This means that the residual size after the appearance of a fault is large and long enough to be detectable.

This chapter provides an overview on different fault diagnosis strategies, with particular attention to the fault detection and isolation (FDI) methods related to the dynamic processes and application examples considered in this book.

For all of the methods considered, it is essential that the technical process can be described by a mathematical model. As there is almost never an exact agreement

Diagnosis and Fault-tolerant Control 2,
coordinated by Vicenç PUIG and Silvio SIMANI. © ISTE Ltd 2021.

between the model used to represent the process and the plant, the model-reality discrepancy is of primary interest.

Hence, the most important issue in model-based fault detection concerns the accuracy of the model describing the behavior of the monitored system. This issue has become a central research theme over recent years, as modeling uncertainty has risen from the impossibility of obtaining complete knowledge and understanding of the monitored process.

The main focus of this chapter is the mathematical description aspects of the process whose faults are to be detected and isolated. The chapter also studies the general structure of the controlled system, its possible fault locations and modes. Residual generation is then identified as an essential problem in model-based FDI, because, if it is not performed correctly, some fault information could be lost. The general framework for the residual generation is also recalled.

Residual generators based on different methods, such as input–output, state and output observers, parity relations and parameter estimations, are just special cases in this general framework. In the following, some commonly used residual generation and evaluation techniques are discussed and their mathematical formulation is presented.

Finally, the chapter presents and summarizes special features and problems regarding the different methods.

1.2. Fault diagnosis tasks

According to the definitions available in the related literature, model-based FDI can be defined as the *detection*, *isolation* and *identification* of faults in a system by using methods that can extract features from measured signals and use *a priori* information on the process available in terms of mathematical models. Faults are, thus, detected by setting fixed or variable thresholds on residual signals generated from the difference between actual measurements and their estimates obtained by using the process model.

A number of residuals can be designed, with each having sensitivity to individual faults occurring in different locations of the system. The analysis of each residual, once the threshold is exceeded, then leads to fault isolation.

Figure 1.1 shows the general model-based FDI system. It comprises two main stages of residual generation and residual evaluation. This structure was first suggested in Chow and Willsky (1980) and now is widely accepted by the fault diagnosis community.

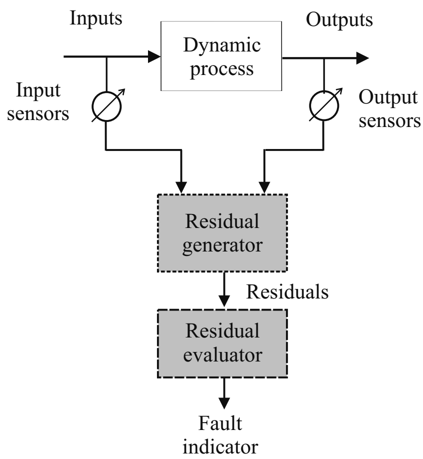

Figure 1.1. *Fault diagnosis module*

The blocks shown in Figure 1.1 perform the following tasks:

1) The residual generation module generates residual signals using the available inputs and outputs from the process under diagnosis. This residual (or fault symptom) should indicate that a fault has occurred. It should normally be zero or close to zero under no fault condition, and significantly different from zero when a fault occurs. This means that the residual is characteristically independent of process inputs and outputs in ideal conditions. Referring to Figure 1.1, this block is called the *residual generator*.

2) The residual evaluation module examines residuals for the likelihood of faults and a decision rule is then applied to determine if any faults have occurred. The *residual evaluator* block in Figure 1.1 may perform a simple threshold test (geometrical methods) on the instantaneous values or moving averages of the residuals. On the other hand, it may consist of statistical methods, for example, generalized likelihood ratio testing or sequential probability ratio testing (Isermann 1997; Willsky 1976; Basseville 1988; Patton *et al*. 2000).

Many works in the field of model-based FDI have focused on the residual generation problem, since the decision-making problem can be considered relatively straightforward if residuals are well designed. In the following, a number of different strategies oriented to solve the model-based residual generation problem have been addressed, with reference to multivariable dynamic processes and technical systems.

The first requirement of the model-based FDI approach consists of providing a mathematical description of the system under investigation that describes the possible fault cases as well.

The general scheme for model-based FDI considered in this chapter is depicted by Figure 1.2. The main components are the *plant* under diagnosis, the *actuators* and the *sensors*, which can be further sub-divided as *input* and *output* sensors, and finally, the *controller*. In the following, the process behavior will be monitored by analyzing its input $u(t)$ and output $y(t)$ measurements and the signals from the controller $u_R(t)$, which are supposedly completely available for FDI purposes. Moreover, as shown in Figure 1.2, the behavior of any controller that drives the system is inherently taken into consideration.

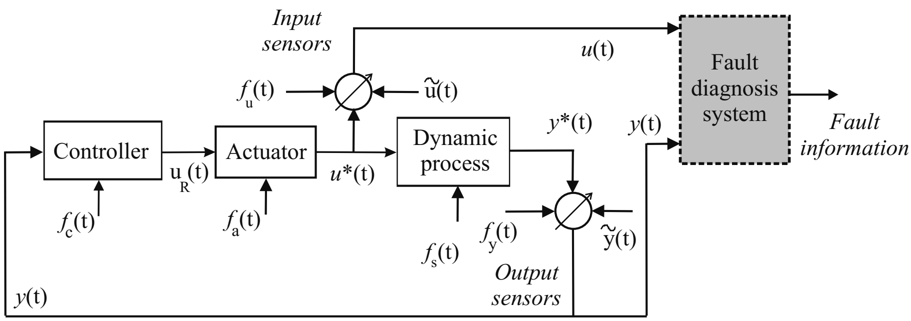

Figure 1.2. *Model-based fault diagnosis strategy*

It is worth noting that, when the signals $u_R(t)$ from the controller or measurements of plant inputs $u(t)$ are not available, the controller plays an important role in the design of the FDI scheme, as a robust controller may reduce the effects of the faults, thus making the fault diagnosis more difficult.

Once the actual process inputs and outputs $u^*(t)$ and $y^*(t)$ (usually not available) are measured by the input and output sensors, FDI theory can be treated as an observation problem of $u(t)$ and $y(t)$. Concerning the occurrence of malfunctions, the *location of faults* and their modeling, the system under diagnosis can be separated into the different parts that can be affected by faults, as illustrated in Figure 1.2. With respect to previous works (see, e.g., Patton *et al.* 1989; Gertler 1998; Patton *et al.* 2000), it is necessary to distinguish between input and output sensors.

Figure 1.2 shows that the input and output signals $u^*(t)$ and $y^*(t)$ are acquired in order to obtain the measurements $u(t)$ and $y(t)$ from the sensors. Figure 1.2 also shows the situation where the controller can be affected by faults, since the monitored process consists of a closed-loop plant. However, because of technological reasons (e.g., the control action is performed by a digital computer), when the actuator is considered as a part or a component of the whole controller device, the former can be treated as subsystem where faults are likelier to occur, while the latter remains free from faults. Under these assumptions, as shown in Figure 1.2, when the monitored process is considered in view of fault location, since input and output measurements are supposed completely available for FDI purposes, the controller behavior in the design of a fault diagnosis scheme can be neglected, as well as the interconnection between the control system and the process.

In general, as shown in Figure 1.2, the actuation signals $u^*(t)$ are assumed to be measurable by neglecting input and output sensor noises. On the other hand, Figure 1.2 represents the situation where the u_R signals are only measured by the input sensors.

Note that the general models for FDI represented in Figure 1.2 can be described in both the time and frequency domains, respectively, which have been widely accepted in the fault diagnosis literature (Patton *et al.* 1989, 2000; Chen and Patton 1999; Gertler 1998). Under these assumptions, the general model-based FDI problem treated here can be performed on the basis of the knowledge of only the measured sequences $u(t)$ and $y(t)$.

As mentioned in Volume 1, frequency domain descriptions are typically applied when the effects of faults, as well as the disturbances, have frequency characteristics which differ from each other and thus, the information in the frequency spectra serves as criteria to distinguish the faults (Ding and Frank 1990; Massoumnia *et al.* 1989). On the other hand, since state–space descriptions provide general and mathematically rigorous tools for system modeling and robust residual generation, for both the deterministic (noise free measurements) and the stochastic case (measurements affected by noises), the model parameters in suitable representations can be obtained by multivariable modeling or identification procedures.

Finally, it is worth noting that most systems to be monitored are actually nonlinear, and the main modeling approaches are recalled here. In fact, there is certainly an increasing interest in the use of nonlinear methods (nonlinear observers, extended Kalman filters, fuzzy-logic methods, and neural networks (NNs)). However, as the feature of system supervision is to monitor the operation and performance of the system with respect to an expected point of operation, linear system methods can be very valid. Deviations from expected behavior can be used to monitor system performance changes, as well as component malfunctions.

1.2.1. *Residual generation task*

This section recalls the general structure of the residual generator for fault diagnosis. The basic methods will be described briefly, while their presentation and application to dynamic processes will be shown in the following chapters.

The residual generator module introduced in Figure 1.1 can be interpreted as illustrated in Figure 1.3 (Basseville 1988).

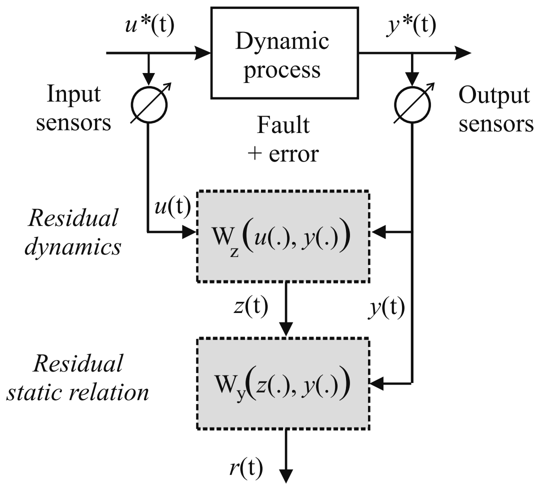

Figure 1.3. *Residual generation strategy*

In the above structure, the auxiliary redundant signal $z(t)$ is generated by the function $W_z\big(u(\cdot), y(\cdot)\big)$ and, together with the measurement $y(t)$, the symptom signal $r(t)$ is computed by means of $W_y\big(z(\cdot), y(\cdot)\big)$.

In the fault-free case, W_y will be almost zero. On the other hand, when a fault occurs in the plant, the residual $r(t)$ will be different from zero.

The simplest residual generator that can be implemented is obtained when the system W_z is a plant identical model $z(t) = W_z\big(u(\cdot)\big)$, or it is an input–output description for the actual plant simply obtained from the system simulator, as described in Volume 1.

In the former case, the measurement $y(t)$ is not required in W_z because it is a process simulator. The signal $z(t)$ represents the simulated output and the residual is directly computed as $r(t) = z(t) - y(t)$. An extension to the model-based residual generation is to replace $W_z\big(u(\cdot)\big)$ by $W_z\,(u(\cdot),\, y(\cdot))$, i.e., an *output estimator* fed by both system inputs and outputs. In such a case, the function W_z generates an estimation of the function of the output.

Therefore, no matter which type of method is used, the residual generation process is nothing but a mapping whose inputs consist of process inputs and outputs.

On the other hand, by considering an input–output approach, Figure 1.4 represents a general structure for all residual generators using the input–output description (Patton and Chen 1991).

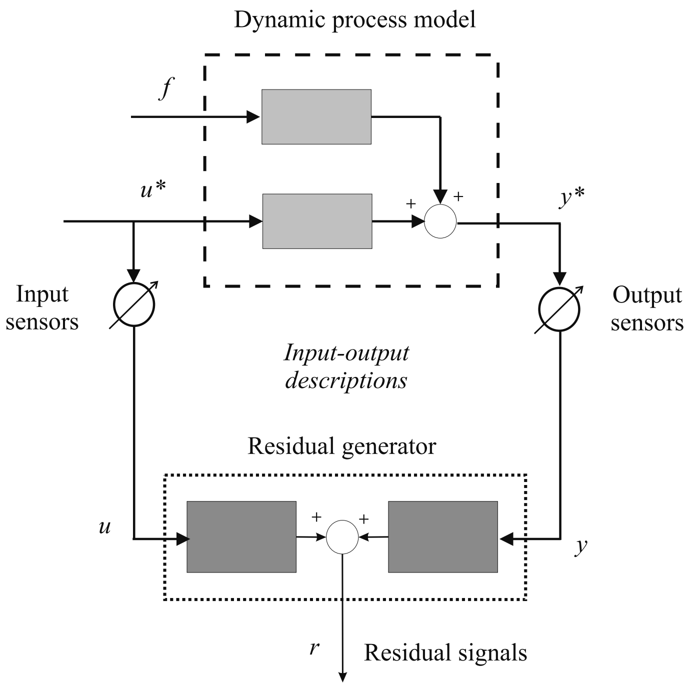

Figure 1.4. *Residual generator input–output form*

According to the definition, the residual function $r(t)$ has to be designed to become almost zero for the fault-free case and significantly different from zero in case of failures. It is worth noting that different residual generators can be obtained by using different parameterizations, as shown in Chen and Patton (1999).

After generating the residual, the simplest and most widely used fault detection is achieved by directly comparing a residual signal $r(t)$ or a residual function $J\left(r(t)\right)$ with a fixed threshold or a threshold function. Note that the fault-free and the faulty conditions depend on $f(t)$, which represents the general fault vector. If the residual exceeds the threshold, a fault may have occurred.

It can be shown that simple geometric tests work especially well with fixed thresholds if the process operates approximately in a steady state and reacts after a relatively large feature, that is, after either a large sudden or a long-lasting gradually increasing fault. On the other hand, adaptive thresholds that depend on plant operating conditions can be exploited, for example when they are expressed as functions of plant inputs (Chen and Patton 1999).

1.2.2. *Residual evaluation task*

When the residual generation stage has been performed, the second step requires the examination of symptoms in order to determine if any faults have occurred.

As suggested above, a decision process may consist of a simple threshold test on the instantaneous values of moving averages of residuals. On the other hand, because of the presence of noise, disturbances and other unknown signals acting upon the monitored system, the decision-making process can exploit statistical methods. In this case, the measured or estimated quantities, such as signals, parameters, state variables or residuals, are usually represented by stochastic variables $r_i(t)$, with the mean value and variance defined as (Willsky 1976):

$$\bar{r}_i = E\{r_i(t)\}; \qquad \bar{\sigma}_i^2 = E\{[r_i(t) - \bar{r}_i]^2\} \tag{1.1}$$

that represent the nominal values for the fault-free process.

Analytic symptoms are then obtained as changes:

$$\Delta r_i = E\{r_i(t) - \bar{r}_i\}; \qquad \Delta\sigma_i = E\{\sigma_i(t) - \bar{\sigma}_i\} \tag{1.2}$$

with reference to the normal values. Usually, the time instant $t > t_f$ represents the unknown instant of the fault occurrence.

In order to separate nominal from faulty behavior, a fixed threshold Δr_{tol} defined as:

$$\Delta r_{tol} = \epsilon\bar{\sigma}_r, \qquad \epsilon \geq 2 \tag{1.3}$$

usually has to be selected. By a proper choice of ϵ, a good compromise has to be made between the detection of small faults (missed fault rates) and false alarm rates.

Another class of methods can be exploited for detecting residual changes due to faults. Therefore, techniques of change detection, for example, as a likelihood-ratio test or Bayes decision, a run–sum test, are commonly used (Isermann 1984; Basseville and Benveniste 1986; Basseville and Nikiforov 1993). Moreover, fuzzy or adaptive thresholds may improve the binary decision (Chen and Patton 1999; Patton *et al*. 2000).

Finally, when several variables change, classification methods can be used. In a multidimensional space, the symptom vector is defined by its components Δr_i, such that Δr belongs to a q-dimensional space and its direction depends on the fault occurrence.

In this case, the process of residual evaluation consists of determining the direction, as well as the distance of Δr from the origin. Geometrical distance methods (Carpenter and Grossberg 1987; Tou and Gonzalez 1974) or artificial NNs (Himmelblau *et al*. 1991; Meneganti *et al*. 1998) can thus be exploited for this purpose.

Hence, the generation and evaluation of analytic symptoms concludes the task of fault detection within the framework of model-based fault diagnosis shown in Figure 1.1.

1.3. Model-based fault diagnosis

The generation of symptoms is the main issue in model-based fault diagnosis. A variety of methods are available in literature for residual generation and this chapter briefly presents some of the most common methods. Most of the residual generation techniques are based on both continuous and discrete system models; however, in this chapter, the attention is mainly focused on continuous-time dynamic models.

The model-based residual generation schemes relying on parity space (relation) methods (Gertler and Singer 1990; Patton and Chen 1991; Gertler and Monajemy 1993; Delmaire *et al*. 1999) and observer-based approaches (Beard 1971; Frank 1993; Frank and Ding 1997; Patton and Chen 1997; Willsky 1976; Basseville 1988) will be considered and summarized (Isermann and Ballé 1997; Patton *et al*. 2000) for their application to dynamic processes and technical systems.

1.3.1. *Parity space relations*

The basic idea of the parity relations approach is to provide a proper check of the parity (consistency) of the measurements acquired from the monitored system. In the early development of fault diagnosis, the parity vector (relation) approach was applied to static or parallel redundancy schemes (Potter and Suman 1977), which may be obtained directly from measurements (hardware redundancy), or from analytical relations (analytical redundancy). A survey of these methods can be found in Ray and Luck (1991). In the case of hardware redundancy, two methods can be exploited to obtain redundant relations. The first requires the use of several sensors with identical or similar functions to measure the same variable. The second approach consists of using dissimilar sensors to measure different variables, but with their outputs being relative to each other. Even if these techniques have been successfully applied for fault diagnosis (Potter and Suman 1977; Daly *et al*. 1979), this section is mainly focused on analytical forms of redundancy.

A straightforward model-based method of fault detection is to take a model of the process G_M and run it in parallel to the process described by G_P, thereby forming an error vector r. The general methodology described here is represented in Figure 1.5.

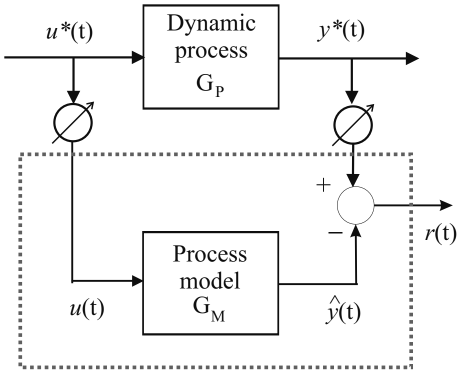

Figure 1.5. *Parity vector approach*

However, as for observers, the model parameters and structure of the monitored process have to be known *a priori*. With reference to Figure 1.3, if $G_M = G_P$, for input f_u and output f_y faults, the r residual has the form:

$$r = G_M\, f_u + f_y \qquad [1.4]$$

In general, the equations (equation [1.4]) that generate the residual signals are called *parity equations* (Gertler 1991) under the assumptions of fault occurrence and exact agreement between process (G_P) and model (G_M). However, within the parity equations, the model parameters are assumed to be known and constant, whereas the parameter estimations can vary the parameters of the process model in order to minimize the residuals. Moreover, for the generation of specific characteristics of the parity vector r and for obtaining fault detection and isolation properties, the residuals can be filtered according to matrix G_f to compute the vector r_f (Patton *et al.* 2000) in the form:

$$r_f = G_f\, r \qquad [1.5]$$

Equations [1.5] and [1.4] can therefore be used to implement and design the residual generation system, in order to meet fault detection and isolation specifications, as well (Gertler 1998). However, for SISO processes only one residual can be generated, and it is therefore not easy to distinguish between different faults.

On the other hand, more freedom in the design of parity equations can be obtained when intermediate signals can be measured for SISO processes (see Figure 1.3) or for Multi-Input Multi-Output (MIMO) systems. As an extension of the parity equation method, the parity relation concept presented here can be generalized

(Chow and Willsky 1984; Lou *et al.* 1986; Patton and Chen 1994b) and then extended to state–space descriptions, as shown in Gertler (1998) for discrete-time models.

In the discrete-time case, the redundancy relations can be exploited to implement a scheme, as shown in Figure 1.6 (Patton and Chen 1994b; Chen and Patton 1999).

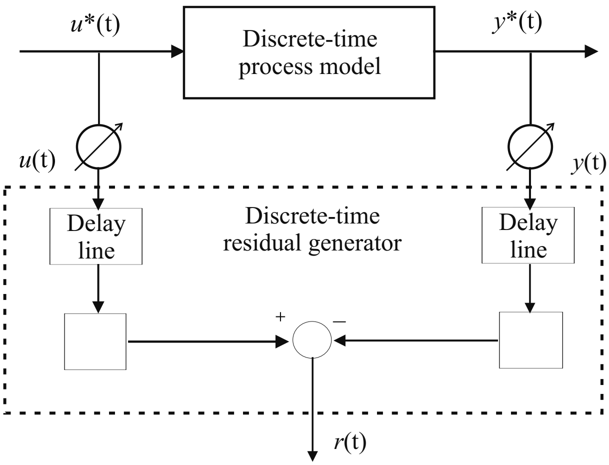

Figure 1.6. *MIMO parity vector*

Finally, because of the previous results, it is clear that some correspondence exists between parity relation and observer-based approaches. This aspect was first pointed out by Massoumnia (1986) and was later demonstrated by Wünnenberg (1990) and Patton *et al.* (1989). The problem was re-examined in detail in Patton and Chen (1994b), and the equivalence under different conditions and in different meanings was discussed. It was shown that the discrete-time parity relation approach is equivalent to the use of a dead-beat observer. This implies that the discrete-time parity relation scheme provides less design flexibility when compared with methods, which are based on observers without any restriction.

A further comparison between observer-based and parity space techniques was proposed (Delmaire *et al.* 1999). Both of the methods were first explored for SISO systems and therefore extended the comparison to MIMO systems. The comparison was performed using discrete-time models. In particular, considering MIMO systems described by estimated input–output discrete-time forms (e.g., Nonlinear ARX [NARX] or Nonlinear Auto Regressive Moving Average eXogenous [NARMAX] models), the parity relations lead to a representation in which parameter redundancy cannot be avoided. To overcome this drawback, Delmaire *et al.* (1999) proposed to use observers designed from identified canonical state–space forms. Moreover, in the case of parameter redundancy, multiple identification of some parameters may occur, leading to inconsistent estimations that may produce inconsistent FDI decisions (Delmaire *et al*. 1999).

1.3.2. *Observer-based approaches*

The basic idea behind the observer-based techniques is to estimate the outputs of the system from the measurements by using either dynamic observers in a deterministic setting or stochastic filters in the noisy environment. The output estimation error (or its weighted value) is therefore used as residual.

It is worth noting that when an observer is exploited for FDI purpose, the estimation of the outputs is necessary, while the estimation of the state vector is usually not needed (Chen and Patton 1999). Moreover, the advantage of using the observer approach is the flexibility in the selection of its gains, which leads to a rich variety of FDI schemes (Frank 1994; Frank and Ding 1997; Chen *et al*. 1996; Liu and Patton 1998).

In order to properly use a (generalized) observer, the dynamic model for the plant under consideration has to be considered:

$$\begin{cases} \dot{x}(t) = f\left(x(t),\, u(t)\right) \\ y(t) = g\left(x(t),\, u(t)\right) \end{cases} \qquad [1.6]$$

being $u(t) \in \Re^r$, $x(t) \in \Re^n$ and $y(t) \in \Re^m$.

Assuming that all of the model parameters are perfectly known, an observer is used to reconstruct the system variables based on the measured inputs and outputs $u(t)$ and $y(t)$:

$$\begin{cases} \dot{z}(t) = h\left(\hat{z}(t),\, u(t),\, y(t)\right) \\ \hat{y}(t) = k\left(\hat{z}(t),\, u(t),\, y(t)\right) \end{cases} \qquad [1.7]$$

The observer scheme described by equation [1.7] is sketched in Figure 1.7.

If the process is influenced by disturbance and faults, by considering the relations of the monitored process model, the overall system includes $v(t)$, which represents the non-measurable disturbance vector at the input, $w(t)$, the non-measurable disturbance vector at the output and $f(t)$ fault signals at the input and output, acting through proper terms. These terms can represent actuator, process, input and output sensor additive faults acting on the considered system.

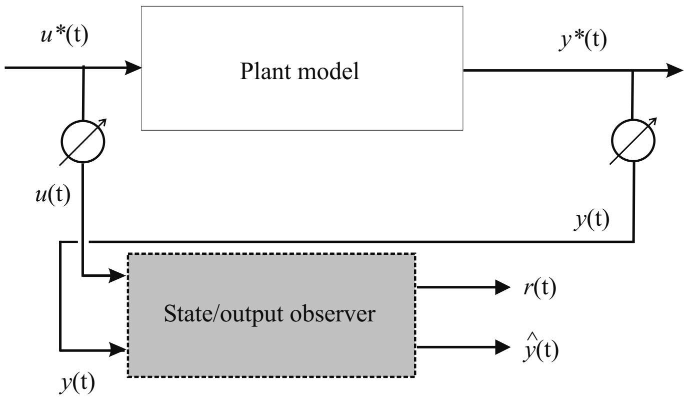

Figure 1.7. *Diagnostic residual observer*

When sudden and permanent faults $f(t)$ occur, the state estimation error $e_x(t)$ will deviate from zero. Moreover, the state estimation error and the output error $e(t)$ show dynamic behavior, which are different for input and output faults. Both $e_x(t)$ and $e(t)$ can be taken as residuals.

In particular, the residual $e(t)$ is the basis for different fault detection methods based on output estimation. For the generation of a residual with special properties, the design of the observer "gain" can be of interest (Chen and Patton 1999; Liu and Patton 1998).

Limiting conditions are the stability and the sensitivity against disturbances $v(t)$ and $w(t)$. If the signals are affected by noise, the extended Kalman filter must be used instead of classical observers (Jazwinski 1970).

The robustness and reliability features of the designed residual generators represent the key point when the proposed solutions have to be applied to real processes.

The parameter changes in the process model can represent both the so-called model-reality mismatch or *multiplicative faults* (Isermann 1997; Patton *et al*. 2000). In the latter case, the changes in the residuals depend on the parameter changes, as well as input and state variable changes. Hence, the influence of parameter changes on the residuals is not as straightforward as in the case of the additive faults $f(t)$. Also, in this case the residuals are only dependent on fault signals $f(t)$ (Patton and Chen 1994a; Chen *et al*. 1996; Gertler 1998; Patton *et al*. 2000).

Finally, a generalization of the techniques above leads to residual generation schemes based on *dedicated observers for MIMO processes* that exploit the following properties of the output observers:

1) *observer excited by one output*: one observer is driven by one sensor output. The other outputs $\hat{y}(t)$ are reconstructed and compared with measured outputs $y(t)$. This allows the detection of single output sensor faults (Clark 1978);

2) *bank of observers, excited by all outputs*: several observers are designed for definite fault signals detected by a hypothesis test (Willsky 1976);

3) *bank of observers, excited by single outputs*: several observers for single sensor outputs are used. The estimated outputs $\hat{y}(t)$ are compared with the measured outputs $y(t)$. This allows the detection of multiple sensor faults (Dedicated Observer Scheme) (Clark 1978);

4) *bank of observers, excited by all outputs except one*: as mentioned before, but each observer is excited by all outputs except one sensor output, which is supervised by a Generalized Observer Scheme (Wünnenberg and Frank 1987; Frank 1993).

1.3.3. *Nonlinear filtering methods*

With reference to the measurement sensors, when the signal to noise ratios $\|u^*(t)\|_2^2/\|\tilde{u}(t)\|_2^2$ and $\|y^*(t)\|_2^2/\|\tilde{y}(t)\|_2^2$ are low, suitable filters must be employed to improve the performance of the FDI system.

For the linear case, with reference to the time-invariant, discrete-time, linear dynamic system, a Kalman filter (KF) can be exploited for the estimation of the process output, as described in Jazwinski (1970).

It can be proved that the KF innovation is a zero-mean white process when all of the assumptions regarding the process model and the statistical characteristics of the noise process affecting the input–output measurements are completely fulfilled. A Riccati equation is thus obtained.

In the presence of a fault on process output, the stochastic properties (mean-value, variance and whiteness, etc.) of the innovation process $e_i(t)$ change abruptly so that the fault detection can be based on these variations, as described in Basseville (1988).

In a similar way to the approach relying on KF, an extended Kalman filter (EKF) can be proposed to solve the FDI fault detection and diagnosis (FDD) problem of fault diagnosis of dynamic processes and technical systems (Kalman 1990). Usually, the methodology is based on joint parameter and state estimation techniques and consists of providing an (optimal) estimate of the fault, which could be used for fault-tolerant control purposes, as described in the following chapters.

Also, in this case the following nonlinear state–space model in the discrete-time framework is considered in the form of equation [1.8]:

$$\begin{cases} x(t+1) = h_i\left(u(t),\, x(t),\, f_i(t),\right) + v(t) \\ y(t) \qquad = g\left(u(t),\, x(t)\right) + w(t) \end{cases} \qquad [1.8]$$

where $h_i(\cdot)$ and $g(\cdot)$ are nonlinear functions and $f_i(t)$ refers to the ith fault function (unknown input) to be diagnosed. The index "i" is used to outline that the estimation of the ith filter is used to estimate the fault $f_i(t)$. The stochastic inputs v and w denote the process and measurement noises, respectively, which are assumed to be uncorrelated white noise processes with covariance matrices:

$$Q(t) = E\left\{v(t)\, v(t)^T\right\},\, R(t) = E\left\{w(t)\, w(t)^T\right\} \qquad [1.9]$$

The initial estimates of the state and covariance matrix are denoted by:

$$\overline{x}_0 = E\left\{x_0\right\}\, P_0 = E\left\{\left(x_0 - \overline{x}_0\right)\, \left(x_0 - \overline{x}_0\right)^T\right\} \qquad [1.10]$$

Following the method proposed in Norgaard *et al.* (2000), the problem of recursively estimating the augmented state vector x can be formulated as a nonlinear filtering problem that minimizes the conditional mean-square error, that is:

$$\hat{x}(t) = \text{argmin} E\left\{\tilde{x}(t)^T\, \tilde{x}(k)|Y^{t-1}\right\} \qquad [1.11]$$

where $\tilde{x}(t) = x(t) - \hat{x}(t)$ represents the state estimation error, while $Y^{t-1} = \left\{y_0,\, y_1,\, \cdots,\, y^{k-1}\right\}$ is a matrix containing the past measurements. The state estimate $\hat{x}(k)$ is equivalent to the conditional mean of the Gaussian probability density function $p\left(x(t)/Y^{(t-1)}\right) \sim \mathcal{N}\left(\hat{x}(t),\, P(t)\right)$ such as:

$$\hat{x}(t) = E\left\{x(t)|Y^{(t-1)}\right\} \qquad [1.12]$$

and where:

$$P(t) = E\left\{\left(x(t) - \hat{x}(t)\right)\, \left(x(t) - \hat{x}(t)\right)^T |Y^{(k-1)}\right\} \qquad [1.13]$$

refers to the state covariance matrix that is used to quantify the uncertainty of the estimate. The estimation algorithm can then be formulated into the following nonlinear observer-based scheme:

$$\begin{cases} \hat{x}(t+1) = f_i\left(u(t),\, \hat{x}(t), f_i(t),\right) + K(t)\, e(k) \\ \hat{y}(t) \qquad = g\left(u(t),\, \hat{x}(t)\right) \end{cases} \qquad [1.14]$$

where $K(t)$ is a non-stationary gain to be computed and $e(t) = y(t) - \hat{y}(t|t-1)$ is the innovation sequence associated with the covariance matrix P_{ee}:

$$P_{ee} = E\left\{(y(t) - \hat{y}(t))\ (y(t) - \hat{y}(t))^T\ |Y^{k-1}\right\} \qquad [1.15]$$

Based on the previous estimate of the state $\hat{x}(t|t)$ with covariance $\hat{P}(t|t)$, the filter computes at a subsequent time-step an optimal estimation of the state $\hat{x}(t+1|k)$ and its covariance matrix $\hat{P}(t+1|k)$ whenever observations become available. This leads to the following update equations:

$$\begin{cases} \hat{x}(t+1) & = \hat{x}(t) + K(t)\, e(t) \\ P(t+1) & = P(t) - K(t)\, P_{ee}(t)\, K^T(t) \end{cases} \qquad [1.16]$$

The expression of $K(t)$ is given in the form of equation [1.17]:

$$K(t) = P_{xy}(t)\, P_{ee}^{-1}(t) \qquad [1.17]$$

where P_{xy} denotes the predicted cross-correlation matrix defined in the following form:

$$P_{xy} = E\left\{(x(t) - \hat{x}(t))\ (y(k) - \hat{y}(k))^T\ |Y^{t-1}\right\} \qquad [1.18]$$

Since the above statistical expectations are generally difficult to obtain, some kind of approximation must be used, like for the EKF case, which exploits a first-order Taylor linearization. However, even if the EKF estimator seems to be adapted, some well-known drawbacks exist in practice, that is, the parameters estimates can converge slower than the state estimates and, in general, only local convergence can be expected. On the basis of the work reported in Norgaard *et al*. (2000), this motivated the use of an approximation of the nonlinear function $h_i(\cdot)$ by means of a multi-dimensional extension of Stirling's interpolation formula (Zolghadri 1996).

Although this method presents some optimality proofs, the key feature remains the *a priori* choice of the covariance matrices Q and R. The matrix Q controls the flexibility of the model, whereas the measurement covariance matrix R controls the flexibility of the measurement equations. In the most practical cases, the optimization of Q and R is done by iteratively testing different values and evaluating the results over a test period.

In practice, this tuning problem is often tackled as an *ad hoc* process involving a very large number of manual trials. In view of this difficulty, it has been chosen to automatically tune these matrices by means of an optimization method (Simani *et al.* 2003). The performance index to be minimized corresponds to the root-mean square of the state estimate errors subjected to positivity constraints of Q and R matrices, that is:

$$J(t) = \left(\frac{1}{N} \sum_{t_0}^{t_f} (\tilde{x}^T \, \Pi \, \tilde{x}) \right)^{\frac{1}{2}} \text{ s.t.} \begin{cases} Q > 0, R > 0 \\ R = diag(r_i) \\ Q = diag(q_i) \end{cases} \qquad [1.19]$$

For convenience, the additional constraints $Q = diag(q_i)$ and $R = diag(r_i)$ are included in the optimization algorithm. Π is a weighting matrix introduced to manage each component of the vector $\tilde{x}$ separately. t_0 and t_f are, respectively, the initial and final discrete time of the tuning interval, and N denotes the number of data points in the tuning interval. Because of the multi-parameter, nonlinear and discrete nature of this optimization problem, Particle Swarm Optimization (PSO) algorithms can be exploited and retained to derive numerical solutions.

1.3.4. *Nonlinear geometric approach strategy*

This section addresses the nonlinear geometric approach (NLGA) to FDI that was proposed in Bonfè *et al.* (2006). The classical NLGA technique is summarized in the following. Moreover, a procedure to obtain suitable NLGA filters for the estimation of the fault affecting the dynamic processes and technical systems that can be used for the active fault tolerant control (AFTC) task is recalled in this book.

The NLGA approach to the nonlinear FDI problem was originally suggested in De Persis and Isidori (2000) and was formally developed in De Persis and Isidori (2001). It consists of finding, by means of a coordinate change in both the state space and in the output space, an observable subsystem which, if possible, is affected by the fault and not affected by disturbance. In this way, necessary and sufficient conditions for the FDI problem to be solvable are given. Finally, a residual generator can be designed on the basis of the model of the observable subsystem. In this work, the complete NLGA strategy with its further extensions and developments is applied to the nonlinear model of the dynamic process under investigation.

In more detail, the NLGA approach considered here requires a nonlinear system model in the form:

$$\begin{cases} \dot{x} = n(x) + g(x)\, c + \ell(x)\, f + p(x)\, d \\ \\ y = h(x) \end{cases} \qquad [1.20]$$

in which $x \in \mathcal{X}$ (an open subset of $\Re^n$) is the state vector, $c(t) \in \Re^{\ell_c}$ is the control input vector, $f(t) \in \Re$ is the fault, $d(t) \in \Re^{\ell_d}$ is the disturbance vector (also embedding the faults that have to be decoupled) and $y \in \Re^m$ is the output vector. $n(x)$, $\ell(x)$, the columns of $g(x)$ and $p(x)$ are smooth vector fields, and $h(x)$ is a smooth map.

Therefore, if P represents the distribution spanned by the column of $p(x)$, the NLGA method can be described by means of the following steps (De Persis and Isidori 2001):

1) determine the minimal conditioned invariant distribution containing P (denoted with Σ_*^P);

2) by using $(\Sigma_*^P)^\perp$, that is, the maximal conditioned invariant codistribution contained in $P^\perp$, determine the largest observability codistribution contained in $P^\perp$, denoted with Ω^*;

3) if $\ell(x) \notin \Omega^*$ continue to the next step, otherwise the fault is not detectable;

4) if the condition of the previous step is satisfied, it can be found that a surjection Ψ_1 and a function Φ_1 fulfilling $\Omega^* \cap \text{span}\{dh\} = \text{span}\{d(\Psi_1 \circ h)\}$ and $\Omega^* = \text{span}\{d(\Phi_1)\}$, respectively. The functions $\Psi(y)$ and $\Phi(x)$ defined as:

$$\Psi(y) = \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \end{pmatrix} = \begin{pmatrix} \Psi_1(y) \\ H_2\, y \end{pmatrix} \qquad \Phi(x) = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{pmatrix} = \begin{pmatrix} \Phi_1(x) \\ H_2\, h(x) \\ \Phi_3(x) \end{pmatrix} \qquad [1.21]$$

are (local) diffeomorphisms, where H_2 is a selection matrix (i.e. a matrix in which any row has all 0 entries but one, which is equal to 1), $\Phi_1(x)$ represents the measured part of the state which is affected by f and not affected by d, and $\Phi_3(x)$ represents the unmeasured part of the state which is affected by f and d.

It is worth noting that Σ_*^P can be computed by means of the following recursive algorithm:

$$\begin{cases} S_0 &= \bar{P} \\ S_{k+1} &= \bar{S} + \sum_{i=0}^{m} \left[g_i, \bar{S}_k \cap \ker\{dh\}\right] \end{cases} \qquad [1.22]$$

where m is the number of inputs, $\bar{S}$ represents the involutive closure of S, $[g, \Delta]$ is the distribution spanned by all vector fields $[g, \tau]$, with $\tau \in \Delta$, and $[g, \tau]$ is the Lie bracket of g, τ. It can be shown that if there exists a $k \geq 0$ such that $S_{k+1} = S_k$, the algorithm [1.22] stops and $\Sigma_*^P = S_k$ (De Persis and Isidori 2001).

Once Σ_*^P has been determined, Ω^* can be obtained by exploiting the following algorithm:

$$\begin{cases} Q_0 &= (\Sigma_*^P)^\perp \cap \text{span}\,\{dh\} \\ Q_{k+1} &= (\Sigma_*^P)^\perp \cap \sum_{i=0}^m \left[L_{g_i} Q_k + \text{span}\,\{dh\}\right] \end{cases} \qquad [1.23]$$

where $L_g\Gamma$ denotes the codistribution spanned by all covector fields $L_g\omega$, with $\omega \in \Gamma$, and $L_g\omega$ the derivative of ω along g.

If there exists an integer k^* such that $Q_{k^*} = Q_{k^*+1}$, Q_{k^*} is indicated as o.c.a. $\left((\Sigma_*^P)^\perp\right)$, where o.c.a. stands for observability codistribution algorithm. It can be shown that $Q_{k^*} = \text{o.c.a.}\left((\Sigma_*^P)^\perp\right)$ represents the maximal observability codistribution contained in $P^\perp$, that is Ω^* (De Persis and Isidori 2001). Therefore, with reference to the model of equation [1.20], when $\ell(x) \notin (\Omega^*)^\perp$, the disturbance d can be decoupled and the fault f is detectable.

In the new (local) coordinate defined previously, the system of equation [1.20] is described by the relations in the following form:

$$\begin{cases} \dot{\bar{x}}_1 = n_1(\bar{x}_1, \bar{x}_2) + g_1(\bar{x}_1, \bar{x}_2)\, c + \ell_1(\bar{x}_1, \bar{x}_2, \bar{x}_3)\, f \\ \dot{\bar{x}}_2 = n_2(\bar{x}_1, \bar{x}_2, \bar{x}_3) + g_2(\bar{x}_1, \bar{x}_2, \bar{x}_3)\, c + \ell_2(\bar{x}_1, \bar{x}_2, \bar{x}_3)\, f + p_2(\bar{x}_1, \bar{x}_2, \bar{x}_3)\, d \\ \dot{\bar{x}}_3 = n_3(\bar{x}_1, \bar{x}_2, \bar{x}_3) + g_3(\bar{x}_1, \bar{x}_2, \bar{x}_3)\, c + \ell_3(\bar{x}_1, \bar{x}_2, \bar{x}_3)\, f + p_3(\bar{x}_1, \bar{x}_2, \bar{x}_3)\, d \\ \bar{y}_1 = h(\bar{x}_1) \\ \bar{y}_2 = \bar{x}_2 \end{cases} \qquad [1.24]$$

with $\ell_1(\bar{x}_1, \bar{x}_2, \bar{x}_3)$ not identically zero.

Denoting $\bar{x}_2$ with $\bar{y}_2$ and considering it as an independent input, the so-called $\bar{x}_1$subsystem written in the following form:

$$\begin{cases} \dot{\bar{x}}_1 = n_1(\bar{x}_1, \bar{y}_2) + g_1(\bar{x}_1, \bar{y}_2)\, c + \ell_1(\bar{x}_1, \bar{y}_2, \bar{x}_3)\, f \\ \bar{y}_1 = h(\bar{x}_1) \end{cases} \qquad [1.25]$$

is affected by the single fault f and decoupled from the disturbance vector d. This subsystem is exploited for the design of the residual generator for the FDI of the fault f.

1.4. Data-driven fault diagnosis

The problem of identifying an unknown system given samples of its behavior is well known (Söderström and Stoica 1987; Ljung 1999) to be ill-posed (Hadamard 1964), as its solution is neither unique nor depends continuously on the given data.

When *a priori* knowledge on the characteristics of the unknown system is available, the identification procedure can be enhanced. This knowledge may act as a set of constraints shaping the space of possible models, so that the identification problem in this new space becomes more tractable. As an example, the regularity of the unknown system can be translated into smoothness constraints of some kind, transforming the identification problem into a minimization problem (Tikhonov and Arsenin 1977; Morozov 1984). This point of view can be successfully applied to estimate algebraic and dynamic affine systems from noisy samples, by assuming certain good properties of the noise and the sampling process (Söderström and Stoica 1987; Ljung 1999). The data-driven methods described in this section start from the results based on the algebraic case, with the purpose of showing the possibility of extending estimation methods to dynamic systems, determining the whole family of models compatible with noisy sequences.

As often happens in new disciplines, systems theory borrowed some tools and viewpoints from existing and well-established fields. Thus, the identification of static and dynamical systems, that is, the determination of models from noisy data, has relied heavily on techniques developed by statisticians, who have traditionally considered it mandatory to associate a unique model to every available set of data, whether contaminated by noise or not. Kalman (1982a, 1982b, 1984) reconsidered this problem, pointing out how the association of a single model to uncertain data is often based on the introduction of additional information, unrelated to the data, that is, of prejudices. While the introduction of such prejudices can be convenient in some practical cases, it is, of course, very important to evaluate the family of solutions that can be found without introducing prejudices or, at least, by introducing only mild ones.

The data-driven approach described in this chapter has started from the algebraic results, with the purpose of investigating the possibility of extending estimation schemes to dynamical systems, determining the whole family of models compatible with noisy sequences. The results obtained differ from the expectations of the authors in that, as it is proved in the following sections, a single model is, in general, compatible with the data. This result is not in contrast with Kalman's result; in fact, in the dynamic case, the additional information necessary to obtain a single model is carried by the correlations established among the samples by the dynamic nature of the generating process.

This section also addresses the problem of the identification of both linear and nonlinear dynamic systems. In the case of nonlinear dynamic systems, the identification will be performed by exploiting parametric nonlinear models, such as affine, neural and fuzzy models.

1.4.1. *Online identification methods*

In most practical cases, the process parameters are not known at all, or they are not known exactly enough. On the other hand, the process can change over time, due to varying working conditions or due to wearing or ageing situations. Then, in these situations, the process model can be determined with data-driven parameter estimation methods, by measuring the input and output signals, $u(t)$ and $y(t)$, if the basic structure of the model is known (Isermann 1997; Patton *et al.* 2000). This strategy can also be considered as a data-driven adaptive approach to FDI, which can be extended to FTC, as shown in the following chapters.

This approach is also based on the assumption that the faults are reflected in the physical system parameters and the basic idea is that the parameters of the actual process are estimated online using well-known parameter estimations methods. The results are thus compared with the parameters of the reference model, obtained initially under fault-free assumptions. Any discrepancy can indicate that a fault may have occurred. In the following, two different techniques can be compared. They exploit different models for describing the input–output behavior of the monitored system.

The first approach relies on the so-called Equation Error (EE) method, that is, the SISO process is described by a discrete-time model of order n that is written in the vector form in the following equation:

$$y(t) = \Psi^T \, \Theta \qquad [1.26]$$

where:

$$\Theta^T = [a_1, \, \ldots a_n, \, b_1, \, \ldots, \, b_n] \qquad [1.27]$$

is the parameter vector and:

$$\Psi^T = [y(t-1), \, \ldots, \, y(t-n), \, u(t-1), \, \ldots, \, u(t-n)] \qquad [1.28]$$

corresponds to the discrete-time data vector. This scheme assumes that the faults and disturbance effects can affect the process parameters, which can thus be used for the change detection task.

According to the representation sketched in Figure 1.8, the equation error $e(t)$ of equation [1.29] is introduced:

$$e(t) = y(t) - \Psi^T \Theta \quad [1.29]$$

or, if:

$$\frac{y(t)}{u(t)} = \frac{B(z)}{A(z)} \quad [1.30]$$

z being the complex variable, and $\frac{B(z)}{A(z)}$ the transfer function of the process, the equation error via Z-transformation becomes:

$$e(t) = \hat{B}(z)u(t) - \hat{A}(z)y(t) \quad [1.31]$$

in which $\hat{A}(z)$ and $\hat{B}(z)$ correspond to the estimates of the parameters of the polynomials in $A(z)$ and $B(z)$.

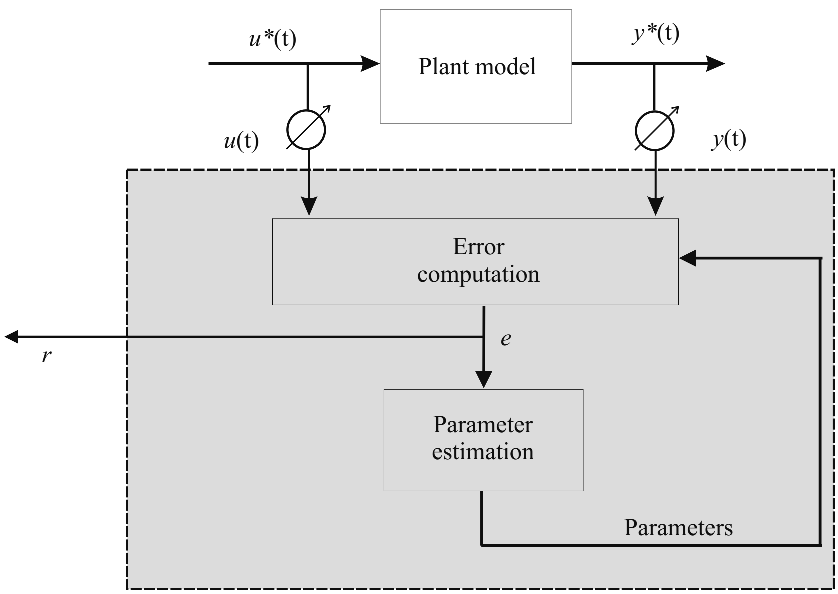

Figure 1.8. *Data-driven fault diagnosis*

The parameter identification approach can be based on the well-known least-squares (LS) estimate:

$$\hat{\Theta} = [\Psi^T \ \Psi]^{-1} \Psi^T y \qquad [1.32]$$

that is, the achieved minimization of the sum of LS is computed according to the following equation:

$$\begin{cases} J(\Theta) & = \sum_t e^2(t) = e^T e \\ \frac{d\ J(\Theta)}{d\ \Theta} & = \qquad 0. \end{cases} \qquad [1.33]$$

As described in Patton *et al*. (2000) and Isermann (1992), the LS estimate can also be expressed in recursive form (RLS) with respect to the estimates at the instant t, with $t = 0,\ 1,\ 2,\ \cdots$:

$$\hat{\Theta}(t+1) = \hat{\Theta}(t) + \gamma(t) \left[y(t+1) - \Psi^T(t+1)\hat{\Theta}(t+1) \right] \qquad [1.34]$$

where:

$$\begin{cases} \gamma(t) & = \frac{1}{\Psi^T(t+1)P(t)\Psi(t+1)+1} P(t)\Psi(t+1) \\ P(t+1) = & \left[I - \gamma(t)\Psi^T(t+1) \right] P(t) \end{cases} \qquad [1.35]$$

For improved estimates, filtering methods can be exploited. In particular, as remarked in section 1.3.3, when the measurements acquired from the process under diagnosis are affected by noise or uncertainty, a filtering technique can be used for the parameter estimation (Jazwinski 1970).

When it is assumed that faults directly affect the process output, an output–error (OE) method can be proposed. Under this assumption, instead of the EE computed in equation [1.29], the OE is described in the form of equation [1.36]:

$$e(t) = y(t) - \hat{y}(\Theta, t) \qquad [1.36]$$

where:

$$\hat{y}(\Theta, z) = \frac{\hat{B}(z)}{\hat{A}(z)}\, u(z) \qquad [1.37]$$

represents the model output, that can also be used, as depicted in Figure 1.8. Unfortunately, in this case, direct calculation of the parameter estimate Θ is not possible, because $e(t)$ is nonlinear in the parameters, as highlighted in equation [1.37].

Therefore, the loss function of equation [1.36] as equation [1.29] has to be minimized by numerical optimization methods. The computational effort is then much larger and online real-time application is, in general, impossible. However, relatively precise parameter estimates may be obtained.

If a fault within the process changes one or several parameters by $\Delta\Theta$, the output signal changes for small deviations according to the description in equation [1.38]:

$$\Delta y(t) = \Psi^T(t)\,\Delta\Theta(t) + \Delta\Psi^T(t)\,\Theta(t) + \Delta\Psi^T(t)\Delta\Theta(t) \qquad [1.38]$$

and the parameter estimator indicates a change $\Delta\Theta$.

Generally, the process parameters Θ depend on physical process coefficients p (such as stiffness, damping factor, and resistance), as represented by equation [1.39]:

$$\Theta = f(p) \qquad [1.39]$$

via nonlinear algebraic relations. If the inversion of the relationship in equation [1.40]:

$$p = f^{-1}(\Theta) \qquad [1.40]$$

exists (Patton *et al*. 2000), the changes Δp of the process coefficients can be calculated. These changes in the coefficients are in many cases, directly related to faults. Thus, although the knowledge of Δp facilitates the fault diagnosis problem, it is not necessary for fault detection only. Parameter estimation can also be applied to nonlinear static process models (Isermann 2005).

1.4.2. *Machine learning approaches to fault diagnosis*

This section recalls data-driven approaches that are based on fuzzy systems and NNs, and are used to implement the fault diagnosis block. In this section, a brief introduction on the general structure of a fault diagnosis system relying on fuzzy systems and NNs is proposed. In particular, their architectures of open-loop NARX systems are reported, since they represent, in combination with proper training algorithms, the exploited solutions for the implementation of the fuzzy systems and NN fault estimators.

With reference to fuzzy system modeling, the design of the fault diagnosis module is achieved by means of Takagi–Sugeno (TS) prototypes. Indeed, the unknown relationships between noisy measurements and faults are provided by fuzzy models, which consist of a number of rules connecting the inputs with the output of the system under investigation, on the basis of a knowledge of its dynamics in form of IF $\Longrightarrow$ THEN relations, processed by fuzzy reasoning (Babuka 2012). In fact, the approximation of nonlinear Multi–Input Single–Output (MISO) systems (but extension to MIMO systems can also be considered) can be achieved by the TS fuzzy reasoning, as reported in Fantuzzi and Rovatti (1996) and Rovatti (1996). According to the TS modeling approach, proposed in Takagi and Sugeno (1985b), the consequents become crisp functions of the input, while the antecedents remain fuzzy propositions, therefore the fuzzy rule takes the form of:

$$\begin{aligned} R_i : IF \quad & (x \text{belongs to the } i\text{-th cluster}) \quad THEN \\ & y_i = a_i^T\, x + b_i \end{aligned} \qquad [1.41]$$

where i indicates the number of rules. The antecedent does not differ from the Mamdani rules, with a combined membership function $\lambda_i(x)$ that takes into account the logical connectives expressed by linguistic propositions. The rule consequent function y_i has a defined structure: it is the instance of parameterized function in the affine linear form, where a_i is a parameter vector and b is a scalar offset, while y_i is the ith rule output. The number of rules is considered equal to the number of clusters n_C used when partitioning the data into regions where local linear relations can be assumed (Babuka 2012). Furthermore, the antecedent of each rule defines the degree of fulfillment for the corresponding consequent model, so that the rule global model can be seen as a fuzzy composition of linear local models.

Thus, the TS inference takes the form of the simple algebraic expression of equation [1.42]:

$$\hat{y} = \frac{\sum_{i=1}^{n_C} \lambda_i(x) \, \left(a_i^T\, x + b_i\right)}{\sum_{i=1}^{n_C} \lambda_i(x)} \qquad [1.42]$$

The estimated output $\hat{y}$ is the weighted average of linear functions of the measured input, where the weights are the combined degree of fulfillment of the system input.

It is worth noting that the nonlinear system under diagnosis can have either a static or a dynamic behavior: in the latter case, the considered model input vector x can contain current, as well as previous samples of the system input or output. In this case, in order to introduce the time dependence into the model of equation [1.41], the consequents are considered as linear AutoRegressive models with eXogenous input (ARX) of order o, in which the regressor vector takes the form of:

$$x(k) = \left[y(k-1),\, \ldots,\, y(k-o),\, u(k),\, \ldots,\, u(k-o)\right]^T \qquad [1.43]$$

where u and y are the actual system input and output vectors, and k is the time step. The affine parameters of the TS model can be grouped in:

$$a_i = [\alpha_1^{(i)}, \ldots, \alpha_o^{(i)}, \delta_1^{(i)}, \ldots, \delta_o^{(i)}]^T \quad [1.44]$$

where the $a^{(i)}$ coefficients are associated with the output samples, and the $\delta^{(i)}$ are associated with the input.

An effective approach to the design of a fuzzy inference system (FIS) as an approximator of a complex nonlinear system begins with the partitioning of the available data into subsets characterized by simpler (linear or affine) behavior. A cluster can be defined as a group of data that are more similar to each other than to members of another cluster. The similarity among data can be expressed in terms of their distance from a particular item, exploited as the cluster prototype. Fuzzy clustering provides an effective tool to obtain a partitioning of data in which the transitions among subsets are smooth, rather than abrupt.

Fuzzy clustering allows an item to belong to several clusters simultaneously, with different degrees of fulfillment, whereas the classic crisp clustering relies on mutual exclusive subsets. Different clustering methods have been proposed in literature, see, for example, the review (Jain *et al.* 1999) or the more recent works (Jun *et al.* 2011; Graaff and Engelbrecht 2012).

Typically, the available data consist of noisy measurements acquired from the system. They are grouped into the data matrix Z, whose columns are the vectors z containing the measurements of a single observation of the system under analysis:

$$Z = \begin{bmatrix} z_{11} & \ldots & z_{1N} \\ \vdots & \ddots & \vdots \\ z_{n1} & \ldots & z_{nN} \end{bmatrix} \quad [1.45]$$

where n is the data dimension and N is the number of available observations.

Most fuzzy clustering algorithms are based on the optimization of the c-means goal function $J\,(Z,\, U,\, V)$ performed as follows:

– Z: the data matrix above is defined;

– $U = [\mu_{ik}]$: the so-called fuzzy partition matrix that contains the values of the membership function for the couple ith measurement/k-th cluster is defined;

– $V = [v_1,\, \ldots,\, v_{n_C}]$ is defined and contains the cluster prototypes that have to be determined and that represent the centers from which the distance of each measurement can be calculated.

The widespread c-means goal function adopted in this work was formulated in Bezdek (2013) in the following form:

$$J(Z, U, V) = \sum_{i=1}^{n_C} \sum_{k=1}^{N} (\mu_{ik})^m D_{ikA}^2 \qquad [1.46]$$

with $m > 1$ weighting exponent, and:

$$D_{ik\,A}^2 = \|z_k - v_i\|_A^2 = (z_k - v_i)^T A(z_k - v_i) \qquad [1.47]$$

that is, a squared inner product distance norm, with $i = 1, \ldots, n_C$ and $k = 1, \ldots, N$. The matrix A determines the cluster shape.

The minimization algorithm exploits a series of Picard iterations consisting of the updating of the cluster prototypes and of the partition matrix, until the stop criterion is met (Babuka 2012).

An important point concerns the determination of the optimal number of clusters n_C, as the clustering algorithm operates on the assumption of a certain number of clusters, regardless of whether they are really present in the data. Once the partition matrix has been estimated, the antecedent degrees of fulfillment are easily derived by interpolation or curve fitting methods.

Then, the design of the FIS assumes the form of an identification problem addressed to the estimation of the consequent parameters a_i and b_i of the TS parameters in a noisy environment. The identification scheme adopted in this work was proposed in Simani *et al*. (1999) and successfully exploited in the approximation of nonlinear functions through the piecewise affine models (Fantuzzi *et al*. 2002). This approach is based on the minimization of the prediction errors of the individual TS local models understood as n_C independent problems. Their solutions rely on the so-called Frisch scheme (Beghelli *et al*. 1990) that is usually exploited in connection with the identification of Errors-In-Variables models (Fantuzzi *et al*. 2002).

Considering a discrete-time MISO system, the noise is supposed to affect the input u, as well as the output y measurements in the form of the additive signals $\tilde{u}$, $\tilde{y}$ on the noise-free unmeasurable quantities u^*, y^*:

$$\begin{cases} u(k) = u^*(k) + \tilde{u}(k) \\ y(k) = y^*(k) + \tilde{y}(k) \end{cases} \qquad [1.48]$$

Thus, considering the ith TS consequent of the type of equation [1.42] and the associated dynamic local ARX model of order o with the regressors grouped into the vector x as in equation [1.43], the acquisition of N_i noisy measurement of input and output samples permits the construction of the ith data matrix $X^{(i)}$ defined as:

$$X^{(i)} = \begin{bmatrix} y(k) & x^T(k) & 1 \\ y(k+1) & x^T(k+1) & 1 \\ \vdots & \vdots & \vdots \\ y(k+N_i-1) & x^T(k+N_i-1) & 1 \end{bmatrix} \quad [1.49]$$

The ith covariance matrix $\Sigma^{(i)}$ from the acquired data can be computed as:

$$\Sigma^{(i)} = X^{(i)^T} X^{(i)} \geq 0 \quad [1.50]$$

that is, a positive-definite matrix consisting of the sum of two terms:

$$\Sigma^{(i)} = \Sigma^{(i)^*} + \bar{\bar{\Sigma}}^{(i)} \quad [1.51]$$

where $\Sigma^{(i)^*}$ refers to the noise-free signals, while $\bar{\bar{\Sigma}}^{(i)}$ is the noise covariance matrix, which depends on the unknown noise variances $\bar{\bar{\sigma}}_u$, $\bar{\bar{\sigma}}_y$ through the expression:

$$\bar{\bar{\Sigma}}^{(i)} = diag\,[\bar{\bar{\sigma}}_y\, I,\, \bar{\bar{\sigma}}_u\, I,\, 0] \quad [1.52]$$

The solution of the above-mentioned identification problem requires the estimation of $\bar{\bar{\sigma}}_u$ and $\bar{\bar{\sigma}}_u$, which can be performed by solving the expression in the form of equation [1.53]:

$$\Sigma^{(i)^*} = \Sigma^{(i)} - \tilde{\Sigma}^{(i)} \quad [1.53]$$

with:

$$\tilde{\Sigma}^{(i)} = diag\,[\tilde{\sigma}_y\, I,\, \tilde{\sigma}_u\, I,\, 0] \quad [1.54]$$

in the variables $\tilde{\sigma}_u$ and $\tilde{\sigma}_y$.

In case all of the assumptions regarding the Frisch scheme (Simani *et al.* 1999) are satisfied, there exists one common point belonging to all of the surfaces $\Gamma^{(i)} = 0$, determined as the root locus of equation [1.53] that represents the actual noise variance values $(\bar{\bar{\sigma}}_u,\, \bar{\bar{\sigma}}_y)$. However, in real cases, the Frisch assumptions are commonly violated, so that a unique solution cannot be obtained. In these situations, the identification aims at finding the nearest point of all of the surfaces.

After the computation of the variances, the covariance noise matrix can be built as in equation [1.52], and the linear parameters in each cluster (therefore in each TS consequent) can be finally determined as a solution of the following expression:

$$\left(\Sigma^{(i)} - \bar{\bar{\Sigma}}^{(i)}\right) a_i = 0 \qquad [1.55]$$

Alongside the fuzzy models, a different data-driven approach, based on NNs, has been proposed in order to implement the fault diagnosis block. In this section, after a brief introduction on the general structure, the properties and the functioning of an NN, as well as the architecture, are recalled. They will be exploited in order to implement the NN fault estimators.

In this work, a set of NN estimators is designed and trained in order to reproduce the behavior of the systems under investigation, thus accomplishing the modeling and identification task. The structure of the ith single neuron (Haykin 2009) is also called *perceptron*. It features a MISO system where the output y_i is computed as a function f of the weighted sum v_i of all the n_i neuron inputs $u_{i,1}, \dots, u_{i,n_i}$, with the associated weights $w_{i,1}, \dots, w_{i,n_i}$. The function f, denominated *activation function*, represents the engine of the neuron, as shown in Figure 1.9.

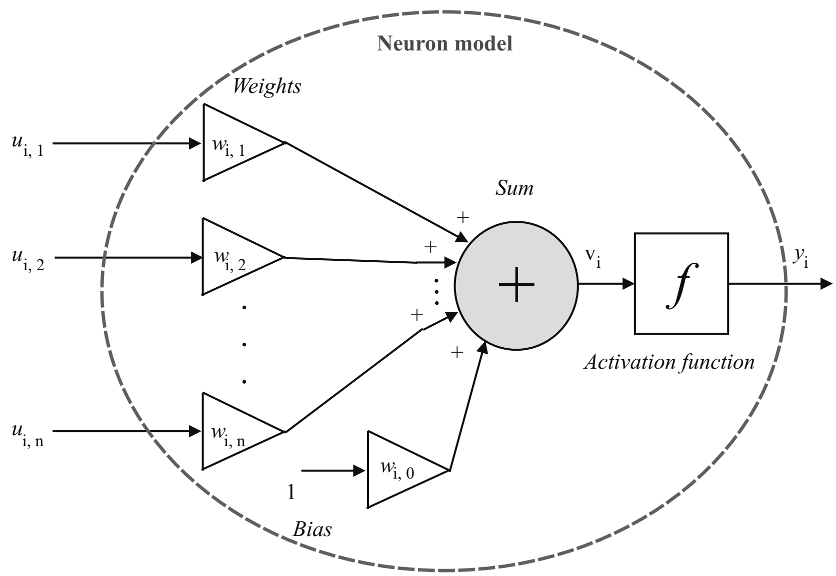

Figure 1.9. *Neuron representation example*

A structural categorization of NNs concerns the way in which their elements are connected with each other (Liu 2012). In a *feed-forward network*, also called *multi-layer perceptron* (MLP), neurons are grouped into unidirectional layers. The first of them, namely the *input* layer, is fed directly by the network inputs, then each successive *hidden* layer takes the inputs from the neurons of the previous layer and transmits the output to the neurons of the next layer, up to the last *output* layer, in which the final network outputs are produced. Therefore, neurons are connected from one layer to the next, but not within the same layer. The only constraint is the number of neurons in the output layer that has to be equal to the number of actual network outputs. On the other hand, *recurrent networks* (Medsker and Jain 1999) are multilayer networks in which the output of some neurons is fed back to neurons belonging to previous layers, thus the information flows forward, as well as backward and allows a dynamic memory inside of the network.

A noteworthy intermediate solution is provided by the MLP with a tapped delay line, which is a feed-forward network whose inputs come from a delay line. This kind of network represents a suitable tool to model, or predict, the evolution of a dynamic system. In particular, the open-loop NARX network belongs to this latter category, as its inputs are delayed samples of the system inputs and outputs. Indeed, if properly trained, a NARX network can estimate the current (or the next) system output on the basis of the acquired past measurements of system inputs and outputs.

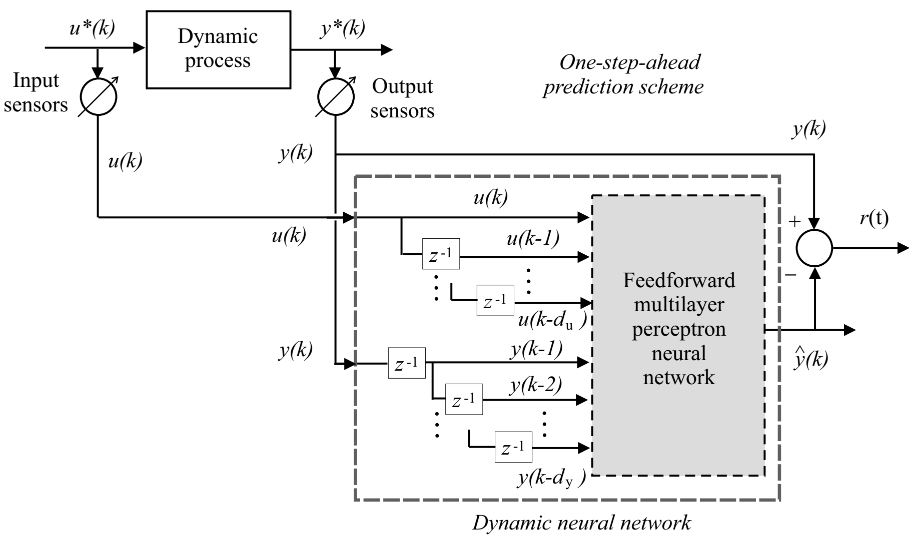

Figure 1.10. *Nonlinear ARX neural network*

Generally speaking, considering a MIMO system, the elaborations of the open-loop NARX network follow the law:

$$\hat{y}(k) = f_{net}\left(u(k),\, \ldots,\, u(k-d_u),\, y(k-1),\, \ldots,\, y(k-d_y)\right) \qquad [1.56]$$

where $\hat{y}$ is the estimation of the system output, u and y are the measured system inputs and outputs, k is the time step, and d_u and d_y are the number of delay of inputs and outputs, respectively. f_{net} is the function realized by the network that depends on the layer architecture, the number of neurons, their weights and their activation functions. The functioning of an open-loop NARX network used as an estimator is depicted in Figure 1.10.

It is worth noting that when only input measurements are available, a NARX network can become a dynamic NN by closing the loop feeding back the network outputs to the inputs, as shown in Figure 1.11.

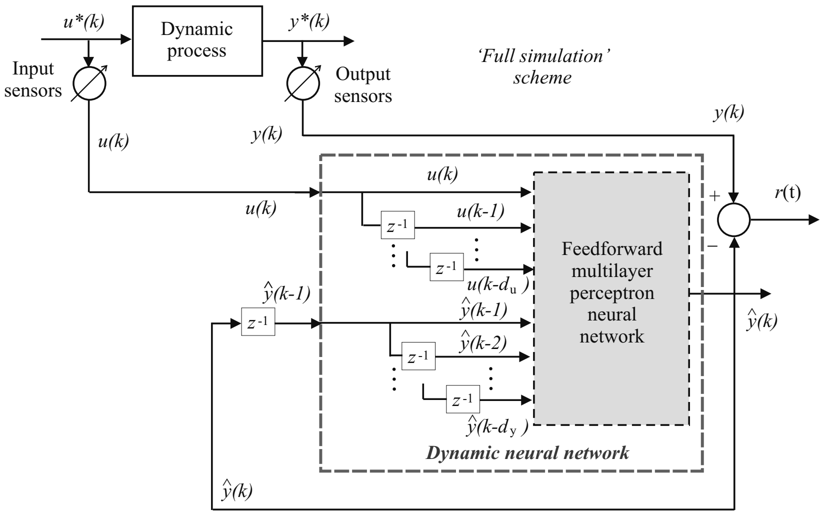

Figure 1.11. *Example of dynamic neural network*

The parameters on which the designer can act concerns the overall architecture (number of neurons, connections between layers), while the value of the weights inside each neuron is derived from the network training.

An NN is a learning system requiring an initial training procedure that adjusts the weights to improve the network performance. When the network task is the estimation of a nonlinear function, the training is performed by presenting a set of examples of proper behavior to the network, consisting of the inputs and the desired outputs (targets) for the relative inputs. Training can be implemented in two different ways as follows:

– **incremental mode**: each couple input–target generates an updating of the network weights;

– **batch mode**: all inputs and targets are applied to the network before the weights are updated.

Although this kind of training requires more memory storage capability, with respect to the incremental mode, it is characterized by a faster convergence and produces smaller errors, thus it will be considered in the following.

The training objective is the minimization of a performance function E, which depends on the weight vector w.

Generally speaking, considering a number P of available example patterns consisting of the input–target pairs (u_p, t_p), with $p = 1, \ldots, P$, with $\hat{y}_p$ represents the output generated by the network fed by u_p. The p-th error vector can be expressed as:

$$e_p = [t_p - \hat{y}_p] = [e_{p,1}, \ldots, e_{p,M}]^T \quad [1.57]$$

with $p = 1, \ldots, P$ and M the number of outputs. Furthermore, the global error vector $\bar{e}$ collects each e_p:

$$\bar{e} = [e_{1,1}, \ldots, e_{1,M}, \ldots, e_{P,1}, \ldots, e_{P,M}]^T \quad [1.58]$$

Consequently, the performance function becomes:

$$E(w) = \frac{1}{P} \sum_{p=1}^{P} (t_i - \hat{y}_i)^2 = \frac{1}{P} \sum_{p=1}^{P} \sum_{m=1}^{M} e_{p,m}^2 \quad [1.59]$$

where the dependence of E by the N parameters grouped in the vector $w = [w_1, \ldots, w_N]^T$ is implicit in the generated output $\hat{y} = \hat{y}(w)$.

Any standard numerical optimization algorithm can be used to update the parameters in order to minimize E. Among these, the most commons are iterative, and make use of characteristic matrices, such as the gradient g (or the Hessian H) of the performance function, or the Jacobian J of the estimation error, defined as:

$$g = \frac{\partial E(w)}{\partial w} = \left[\frac{\partial E}{\partial w_1}, \ldots, \frac{\partial E}{\partial w_N} \right]^T \quad [1.60]$$

$$H = \begin{bmatrix} \frac{\partial^2 E}{\partial w_1^2} & \cdots & \frac{\partial^2 E}{\partial w_1 w_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 E}{\partial w_N w_1} & \cdots & \frac{\partial^2 E}{\partial w_N^2} \end{bmatrix} \quad [1.61]$$

$$J = \begin{bmatrix} \frac{\partial^e_{1,1}}{\partial w_1} & \cdots & \frac{\partial^e_{1,1}}{\partial w_N} \\ \frac{\partial^e_{1,2}}{\partial w_1} & \cdots & \frac{\partial^e_{1,2}}{\partial w_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial^e_{P,M}}{\partial w_1} & \cdots & \frac{\partial^e_{P,M}}{\partial w_N} \end{bmatrix} \quad [1.62]$$

The successive iterations of these algorithms consist of the updating of the parameters and the calculation of the new value of the performance function, until a stop criterion is met. The updating rules of the most common optimization algorithms (i.e. the gradient descent, the Newton, the Gauss–Newton and the Levenberg–Marquardt algorithm) are reported in Table 1.1.

Table 1.1 summarizes the parameters of the updating rules of the most common optimization algorithm, aimed at the minimization of the performance function E. k is the iteration index, α is the learning rate and μ is the combination coefficient.

Optimization method	Adaptation law
Gradient descent	$w_{k+1} = w_k - \alpha\, g_k$
Newton–Raphson	$w_{k+1} = w_k - H_k^{-1}\, g_k$
Gauss–Newton	$w_{k+1} = w_k - \left(J_k^T\, J_k\right)^{-1}\, J_k\, \bar{e}_k$
Levenberg–Marquardt	$w_{k+1} = w_k - \left(J_k^T\, J_k + \mu\, I\right)^{-1}\, J_k\, \bar{e}_k$

Table 1.1. *Training algorithm examples*

It can be demonstrated that the gradient descent algorithm, for a sufficiently small learning rate α value, is asymptotically convergent: around the solution g, the gradient is close to zero and the weights do not meaningfully change. Otherwise, the Newton and the Gauss–Newton algorithms provide a faster convergence, but they both involve the computation of the inverse of a matrix that may not be invertible, causing instability in the procedure. Moreover, the Hessian matrix entails a burdensome computational effort, as it contains the second-order derivative terms.

The *Levenberg–Marquardt* algorithm, originally proposed in Marquardt (1963), introduces an approximation of the Hessian matrix as $H \approx J^T\, J + \mu\, I$, where the first term of the sum is the Jacobian approximation (also exploited in Gauss–Newton) and the second term, driven by the combination coefficient $\mu > 0$, ensures the invertibility of the resulting matrix. Therefore, the Levenberg–Marquardt algorithm provides both a fast and a stable convergence and it represents a suitable tool to train an NN. Indeed, the NN fault estimator blocks have been trained exploiting this method.

The training of an NN based on the Levenberg–Marquardt algorithm, as explained in Hagan and Menhaj (1994), uses a technique denominated *back-propagation* training in order to compute the Jacobian matrix for the updating rule. Its name refers to the backward processing that starts from the output layer of the network toward the first layer, after a previous forward computation of neuron outputs.

1.5. Model-based and data-driven integrated fault diagnosis

Several FDI techniques have been developed and their application shows different properties with respect to the diagnosis of different faults in a process. In order to achieve a reliable FDI technique, a good solution consists of a proper integration of several methods that take advantages of the different procedures (Isermann 1994; Isermann and Ballé 1997). Furthermore, a comprehensive approach to fault diagnosis should exploit a knowledge-based treatment of all available analytical and heuristic information. This successful approach can be performed by an integrated method to knowledge-based fault diagnosis.

Regarding fuzzy logic for residual generation, as stated in section 1.2.1, model-based FDI consists of two stages, residual generation and decision making. The first block is exploited to generate residuals by means of the available inputs and outputs from the monitored system.

For the first step, classical fault diagnosis model-based methods can exploit the state–space of input–output dynamic models of the process under investigation. Within this framework, faults are supposed to appear as changes on the system state or output caused by malfunctions of the components, as well as of the sensors. Such fault indices are often monitored using estimation techniques. The main problem with these techniques is that the precision of the process model affects the accuracy of the detection and isolation system, as well as the diagnostic sensibility. On the other hand, general technical systems are nonlinear, as highlighted in this book, and may not be modeled by using a single model for all operating conditions.

Since a mathematical model is a description of system behavior, accurate modeling for a complex nonlinear system is very difficult to achieve in practice. Sometimes, for some nonlinear systems, it can be impossible to describe them by analytical equations. Moreover, sometimes the system structure or parameters are not precisely known and if the diagnosis has to be based primarily on heuristic information, no qualitative model can be set up. Because of these assumptions, fuzzy system theory seems to be a natural tool to handle complicated and uncertain conditions (Babuška 1998). Therefore, instead of exploiting complicated nonlinear models obtained by modeling techniques, it is also possible to describe the behavior of the nonlinear dynamic process under investigation by a collection of local affine fuzzy and non-fuzzy models (Leontaritis and Billings 1985a, 1985b; Takagi and Sugeno 1985a), whose parameters are obtained by identification procedures.

The second stage of model-based FDI consists of a logic decision process that transforms residual signal information (quantitative knowledge) into qualitative statements (faulty or normal working conditions). Therefore, the problem of decision making can be treated in a novel way by means of fuzzy logic.

Noise contamination and uncertainty affect the residuals, even in fault-free conditions, so that they fluctuate and become unequal to zero. This common situation, which may hide the fault effects, can be handled by means of the fuzzy logic framework. The interesting feature of fuzzy logic is that it represents a powerful tool for describing vague and imprecise facts and is therefore suited for applications where complete information about the fault and the system is not available to the designer.

Even if much effort has been made to decrease the uncertainty associated with quantitative residual generation, it is impossible to fully eliminate the effect of uncertainty. On the basis of this limitation, the residual evaluation problem consists of making the correct decision with respect to uncertain information. Fuzzy logic can be a suitable tool for this task. For instance, a lot of processes can be managed heuristically by humans, since an analytical description is impossible to use. Fuzzy logic can express expert knowledge in the form of a rule-based knowledge format. The introduction of fuzzy logic can thus improve decision making in order to provide reliable FDI methods, which are applicable for real technical systems.

It should finally be pointed out how the fuzzy approach in FDI can solve the problem at two levels: first, fuzzy descriptions are used to generate symptoms, and then the fault detection and isolation are achieved by again using fuzzy logic (Dexter and Benouarets 1997; Isermann 1998).

On the other hand, with reference to NNs in fault diagnosis, quantitative model-based fault diagnosis generates symptoms on the basis of the analytical knowledge of the process under investigation. In most cases, however, this does not provide enough information to perform an efficient FDI, that is, to indicate the location and the mode of the fault.

A typical integrated fault diagnosis system uses both analytical and heuristic knowledge of the monitored system. The knowledge can be processed in terms of residual generation (analytical knowledge) and feature extraction (heuristic knowledge). The processed knowledge is then provided to an inference mechanism, which can comprise residual evaluation, symptom observation and *pattern* *recognition*. In particular, when the process model is only known to a certain extent of precision, the pattern recognition method can provide a convenient approach to solve the fault identification problem, that is, to determine the size of the fault (Himmelblau 1978; Pau 1981).

NNs have been used successfully in pattern recognition, as well as system identification, and they have also been proposed as a possible technique for fault diagnosis. NNs can handle nonlinear behavior and partially known process because they learn the diagnostic requirements by means of the information of the training data. NNs are noise tolerant and their ability to generalize the knowledge, as well as to adapt during use, are extremely interesting properties (Hoskins and Himmelblau 1988; Dietz *et al*. 1989; Venkatasubramanian and Chan 1989; McDuff and Simpson 1990; Chen *et al*. 1990).

Some example processes were considered, in which FDI was performed by an NN using input and output measurements. In these works, the NN is trained to identify the fault from measurement patterns, however the classification of individual measurement patterns is not always unique in dynamic situations; therefore, the straightforward use of NNs in fault diagnosis of nonlinear dynamic processes is not practical, and other approaches should be investigated. An NN could be exploited in order to find a dynamic model of the monitored system or connections from faults to residuals. In the latter case, the NN is used as a pattern classifier or nonlinear function approximator. In fact, artificial NNs are capable of approximating a large class of functions for fault diagnosis of an industrial plant.

Under these considerations, in this chapter, the identification of fuzzy and non-fuzzy models for the system under diagnosis, as well as the application of an NN as a function approximator will be shown. Quantitative and qualitative approaches have a lot of complementary characteristics, which can be suitably combined together to exploit their advantages and increase the robustness of quantitative techniques. The suggested combination can also minimize the disadvantages of the two procedures; in particular, it is important that partial knowledge deriving from qualitative reasoning is reduced by quantitative methods. Hence, the main aim of further research on model-based fault diagnosis consists of finding a way to properly combine these two approaches together to provide highly reliable diagnostic information.

After these remarks, it is worth highlighting the use of neuro-fuzzy (NF) approaches to FDI. The identification of multivariable processes can be interpreted as a problem of approximation to an input–output mapping. The mathematical model used in traditional methods is sensitive to modeling errors, parameter variation, noise and disturbance (Chen and Patton 1999; Patton *et al*. 2000). Process modeling has limitations, especially when the system is complex and uncertain and the data are ambiguous and not rich in information.

As previously stated, NNs are known to approximate any nonlinear even dynamic function, given suitable weighting factors and architecture. Moreover, online training makes it possible to change the FDI system easily in cases where changes are made in the physical process or the control system. NNs can generalize when presented with inputs not appearing in the training data and make intelligent decisions in cases of noisy or corrupted data. They are also readily applicable to multivariable systems and have a highly parallel structure, which is expected to achieve a higher degree of fault tolerance. An NN can operate simultaneously on qualitative and quantitative data. NNs can be very useful when no mathematical model of the system is available, that is, analytical models cannot be applied.

It is clear that almost all of the physical processes are dynamic in nature. Combining dynamic elements such as filters and delays yield a powerful modeling technique. But the NN operates as a "black box" with no qualitative/quantitative information available of the model it represents. Usually, engineers and operators want to visualize how the system is working and what rules govern its operation. There is also ambiguity about the performance of the NN in case of unexpected situations (Korbicz *et al*. 1999).

Fuzzy logic systems, on the other hand, have the ability to mimic the sensing, generalizing, processing, operating and learning abilities of a human operator. They offer a linguistic model of the system dynamics that can be easily understood by certain rules. They also have inherent abilities to deal with imprecise or noisy data. Fuzzy logic can be used with NNs (Chiang *et al*. 2001). A fuzzy neuron has the same basic structure as an artificial neuron, except that some or all of its components and parameters may be described through fuzzy logic. A fuzzy NN is built on fuzzy neurons or on standard neurons dealing with fuzzy data. A fuzzy NN is a connectionist model for the implementation and inference of fuzzy rules. There are many different ways to fuzzify an artificial neuron, which results in a variety of fuzzy neurons and fuzzy networks (Chiang *et al*. 2001; Nelles 2001).

Different NF structures can therefore be designed to combine the advantages of both NNs and fuzzy logic (Patton *et al*. 1999; Calado *et al*. 2001). These structures have been successfully applied to a wide range of applications, from industrial processes to financial systems, because of the ease of rule base design, linguistic modeling, application to complex and uncertain systems, inherent nonlinear nature, learning abilities, parallel processing and fault-tolerance abilities (Wu and Harris 1996; Ayoubi 1995). However, successful implementation depends heavily on prior knowledge of the system and the training data. There are three common methods of combining NNs with fuzzy logic.

1) fuzzification of the inputs or outputs of the NNs;

2) fuzzification of the interconnections of conventional NNs;

3) using NNs in fuzzy models where neurons provide the necessary membership functions and rule base.

All of the NF modeling structures combine, in a single framework, both numerical and symbolic knowledge about the process. Automatic linguistic rule extraction is a useful aspect of NF, especially when little or no prior knowledge about the process is available (Brown and Harris 1994; Jang and Sun 1995). For example, an NF model of a nonlinear dynamical system can be identified from the empirical data. This modeling approach can give us some insight about the nonlinearity and dynamical properties of the system.

The most common NF systems are based on two types of fuzzy models (Takagi and Sugeno 1985a; Sugeno and Kang 1988; Mamdani 1976; Mamdani and Assilian 1995) combined with NN learning algorithms. TS models use local linear models in the consequents, which are easier to interpret and can be used for control and fault diagnosis (Füssel *et al.* 1997; Isermann and Ballé 1997). Mamdani models use fuzzy sets or rules as consequents and therefore give a more qualitative description. The B-spline NN (with triangular basis functions) is the simplest of all of the Mamdani NF structures, but the large consequent rule set means that the method is not easy to use due to low transparency.

Many NF structures have been successfully applied to a wide range of applications from industrial processes to financial systems, because of the ease of rule based design, linguistic modeling, application to complex and uncertain systems, inherent nonlinear nature, learning abilities, parallel processing and fault-tolerance abilities. However, successful implementation depends heavily on prior knowledge of the system and the empirical data (Ayoubi 1995).

NF networks, by their intrinsic nature, can handle a limited number of inputs and can usually be identified in a not very transparent way from the empirical data. Here, transparency corresponds to a more meaningful description of the process, that is, less rules with appropriate membership functions. In adaptive neuro-fuzzy inference systems (ANFIS) (Jang 1993; Jang and Sun 1995), a fixed structure with grid partition is used. Antecedent and consequent parameters are identified by a combination of LS estimates and gradient-based methods, the so-called *hybrid learning rule*. This method is fast and easy to implement for low-dimensional input spaces. It is more prone to losing the transparency and the local model accuracy because of the use of the error back-propagation, that is, a global and not locally nonlinear optimization procedure. One possible method to overcome this problem can be to find the antecedents and rules separately, for example, by clustering and constraining the antecedents, and then applying optimization.

Hierarchical NF networks can be used to overcome the dimensionality problem by decomposing the system into a series of MISO and/or SISO systems called *hierarchical systems* (Tachibana and Furuhashi 1994). The local rules use subsets of input spaces and are activated by higher level rules. The criteria on which to build an NF model are based on the requirements for fault diagnosis and the system characteristics. The function of the NF model in the FDI scheme is also important, that is, pre-processing data, identification (residual generation) or classification (decision making/fault isolation). For example, an NF model with high approximation capability and disturbance rejection is needed for identification so that the residuals are more accurate, whereas, in the classification stage, an NF network with more transparency is required.

In the remainder of this section, the problem of structure identification for NF models is briefly addressed. For complexity reduction and transparency, structure identification methods can be applied to find appropriate input partition, rules and membership functions (MFs). Methods like Evolutionary Algorithms (EA), Classification and Regression Trees (CART) (Jang 1994), clustering and unsupervised NN (e.g. the Kohonen feature maps) can be used. Once the structure is determined, that is, the rules and input membership functions, the consequent parameters can be identified by optimization techniques like the LS estimation. The product space clustering approach can be used (Babuška 1998) for structure identification of TS and Mamdani fuzzy models. For a MISO nonlinear dynamic system with p inputs, the product space $X \times Y \subset \Re^{p+1}$ is divided into subspaces in which linear models can approximate the nonlinear system. The LOcally LInear MOdel Tree (LOLIMOT) algorithm, developed in Nelles and Isermann (1996), can be used to identify a fuzzy model with dynamic linear models as consequent. When using such structure identification techniques, a major issue is the sensitivity to uneven distribution of data. For example, in most clustering algorithms, more clusters are created in regions with more data. A possible solution to this problem may be to initialize the algorithm with a large number of clusters.

Figure 1.12 describes an FDI scheme, including different residual generation and evaluation strategies. Several models are constructed to identify the faulty and fault-free behavior of the system under diagnosis.

$$r_i(t) = f\left(u(t), \ldots, u(t-n), y(t), \ldots, y(t-n)\right) \qquad [1.63]$$

with $i = 1, \ldots, m$. Each residual $r_i(t)$ in equation [1.63] is ideally sensitive to one particular fault in the system. These residual functions can be obtained by both data-driven and model-based approaches, as described in the previous sections. In practice, however, as a consequence of noise and disturbances, residuals can be sensitive to more than one fault. To take into account the sensitivity of residuals to various faults and noise, in the scheme of Figure 1.12, an NF classifier is exploited. In this case, an NF network is used that processes the residuals to indicate the fault. This NF model is constructed with following set of rules:

If r_1 is small $\ldots r_j$ is large, r_m is small **then** fault$_r$ is large. [1.64]

As an alternative to geometric methods for residual evaluation, equation [1.65] employs a fuzzy threshold δ_ν, in order to take into account the imprecision of the residual generator at different regions in the input space.

$$\delta_\nu(u) = \frac{\sum_{i=1}^{C} \delta_i \, \eta_i(u)}{\sum_{i=1}^{C} \eta_i(u)} \qquad [1.65]$$

where C is the total number of regions (or clusters) with different sensitivity to faults and a multidimensional fuzzy set η_i defines the fuzzy boundary of the ith region. This approach depends heavily on the availability of the faulty and fault-free data and it is more difficult to isolate faults that appear in the dynamics.

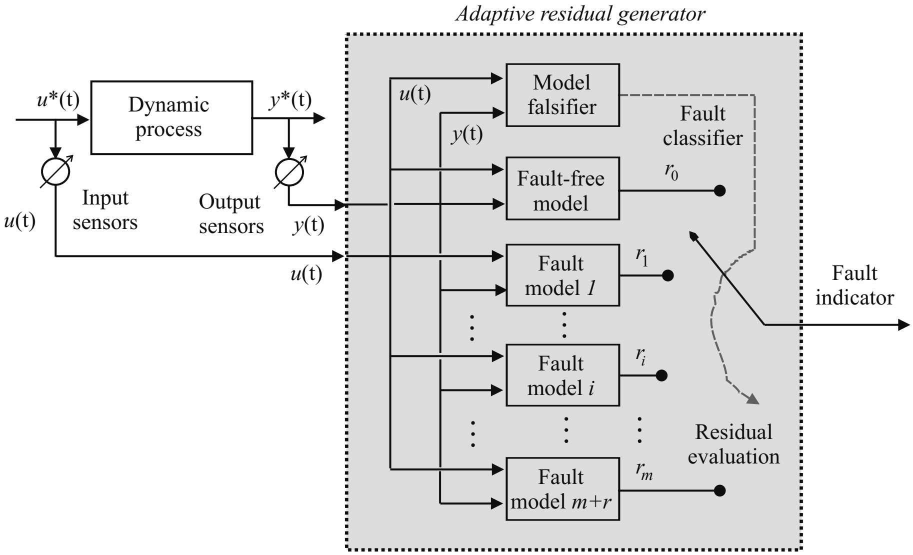

Figure 1.12. *Fault diagnosis approach integration*

Residuals r_i can also be generated by nonlinear dynamic prototypes of the plant that approximate the nonlinear process under diagnosis. These models can be obtained by means of model-based or data-driven techniques, as described in the previous sections. In the case of fuzzy descriptions, they can be derived via product space clustering (Babuška 1998) or the LOLIMOT algorithms (Nelles and Isermann 1996). Each local model is a linear approximation of the process in a subspace and the selection of the local model is fuzzy. The output of such a model can be described by:

$$y(t) = \frac{\sum_{i=1}^{C} \mu_i(x(t))\, y_i(t)}{\sum_{i=1}^{C} \mu_i(x(t))} \qquad [1.66]$$

where $x(t)$ is a suitable combination of the input and output signals, while $y_i(t)$ is the ith output of the local linear (or affine) model given by:

$$y_i(t) = \sum_{k=1}^{n} b_{i,k}\, u(t-k) + \sum_{k=1}^{n} a_{i,k}\, y(t-k) + c_i \qquad [1.67]$$

with $a_{i,k}$, $b_{i,k}$ and c_i as the parameters of the ith model, and $x(t)$ as the subspace defining the operating point, μ_i is the degree to which the ith local model is valid at this operating point.

From $a_{i,k}$, $b_{i,k}$ and c_i, physical parameters such as time constants, static gains, and offsets (Füssel *et al.* 1997) can be extracted for each operating point and can be compared with the parameters estimated online, as mentioned in section 1.4.1. This approach heavily depends on the accuracy of the nonlinear dynamic model described above. Also, the output error should be minimum when operated in parallel to the system. Moreover, this method requires that there is sufficient excitation at each operating point for online estimation of parameters. This residual generation scheme for FDI is sketched in Figure 1.13.

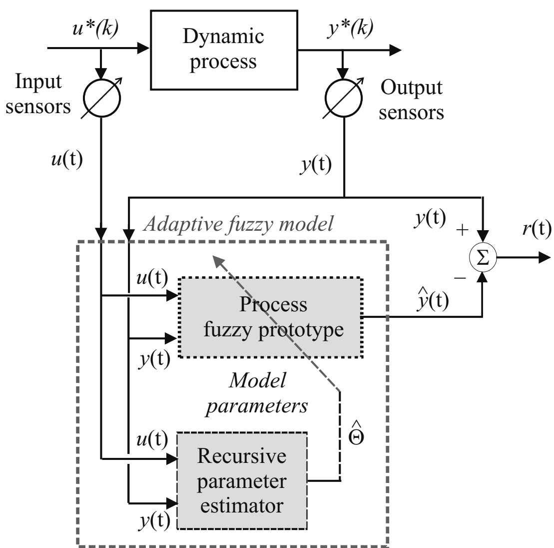

Figure 1.13. *Online estimation for fault diagnosis*

1.6. Robust fault diagnosis problem

Although the analytical redundancy method for residual generation has been recognized as an effective technique for detecting and isolating faults, the critical problem of unavoidable modeling uncertainty has not been fully solved. The main problem regarding the reliability of FDI schemes is the modeling uncertainty, which is due, for example, to process noise, parameter variations and nonlinearities.

On the other hand, all model-based methods use a model of the monitored system to produce the symptom generator. If the system is not complex and can be described accurately by the mathematical model, FDI is directly performed by using a simple geometrical analysis of residuals. In real industrial systems, however, the modeling uncertainty is unavoidable. The design of an effective and reliable FDI scheme for residual generation should take into account the modeling uncertainty with respect to the sensitivity of the faults. Therefore, the task of the design of an FDI system is thus to generate residuals that are *robust* and *reliable* (Frank 1994; Frank and Ding 1997; Patton and Chen 1994b). Several papers addressed this problem. For example, optimal robust parity relations were proposed (Chow and Willsky 1984; Chung and Speyer 1998; Speyer 1999; Lou *et al*. 1986) and the threshold selector concept was introduced (Emami-Naeini *et al*. 1988). Robust FDI using the disturbance decoupling technique was also used (Patton and Chen 1994b; Chen *et al*. 1996). In Patton and Chen (1994b) and Chen *et al*. (1996), this approach represents an interesting contrast to the method proposed in Chow and Willsky (1984), which seems to minimize the modeling uncertainty over several points of operation. In Patton and Chen (1994b) and Chen *et al*. (1996), this problem was solved by estimating the optimum unknown input distribution matrix over a range of operating points, and exploiting the eigenstructure assignment approach (Patton and Chen 1994b; Chen and Patton 1999).

The model-based FDI technique requires a high accuracy mathematical description of the monitored system. The better the model represents the dynamic behavior of the system, the better the FDI precision will be. If an FDI method can be developed, which is insensitive to modeling uncertainty, a very accurate model is not necessarily needed. All uncertainties can be summarized as disturbances acting on the system. Although the disturbance vector is unknown, its distribution matrix can be obtained by using an identification procedure. Under this assumption, the "disturbance decoupling" principle can be exploited to design a robust FDI scheme.

In order to take into account the robustness and reliability feature in the design of the fault diagnosis solution, the model of the monitored system should include disturbance vectors, the known or unknown input–output distribution terms, the parameter errors or variations representing modeling errors and modeling uncertainty (Patton and Chen 1993).

With reference to the residual generator function, these model-reality mismatch effects are present. Both faults and modeling uncertainty (disturbance and modeling error) affect the residual, and hence, discrimination between these two effects is difficult.

The principle of disturbance decoupling for robust and reliable fault diagnosis requires that the residual generator is not affected by uncertainty and input effects. This property can be achieved by using generalized observer schemes (Chen *et al.* 1996; Frank *et al.* 2000), optimal (robust) parity relations (Chow and Willsky 1984; Frank *et al.* 2000) or specific design approaches (Liu and Patton 1998; Patton and Chen 2000; Duan *et al.* 2002).

Hence, for disturbance decoupling approaches in fault diagnosis, the aim is to completely eliminate the disturbance effect from the residual. However, the complete elimination of disturbance effects may not be possible due to the lack of degree of freedom. Moreover, it may be problematic, in some cases, because the fault effect may also be eliminated. Hence, an appropriate criterion for robust residual design should take into account the effects of both modeling error and faults. There is a trade-off between sensitivity to faults and robustness to modeling uncertainty and hence, robust residual generation can be considered as a *multi-objective optimization problem* (Chen and Patton 1999). It consists of the maximization of fault effects and the minimization of uncertainty effects. Therefore, the approach to the design of optimal residuals requires the satisfaction of a set of objectives. These objectives are essential for achieving robust diagnosis of incipient faults. If such joint optimization problems, which can be also expressed in suitable domains, were reformulated for satisfying a set of inequalities on the performance indices, large-scale optimization strategies (Davis 1991; Boyd *et al.* 1994) can be successfully exploited to search for the optimal solution (Chen and Patton 1999, 2001).

Disturbance decoupling can also be achieved using proper design techniques. As an example, the robust fault detection problem can be managed by using the standard H_∞ filtering formulation (Hou and Patton 1996; Frank and Ding 1997). With this method, the minimization of the disturbance effect on the residual is formulated as a standard H_∞ filtering problem (Chen and Patton 2000; Frank *et al.* 2000). On the other hand, the so-called H_∞/H_- approach can be also exploited (Chen and Patton 2000).

Among the many ways for eliminating or minimizing disturbance and modeling error effects on the residual and hence for achieving robustness in FDI (Patton *et al.* 2000), the H_∞ optimization is a robust design method, with the original motivation firmly rooted in the consideration of various uncertainties, especially the modeling errors. It is reasonable to seek an application of this technique in the robust design of FDI systems for dynamic processes and technical systems. Therefore, the H_∞ optimization method can be successfully exploited for robust residual generation of FDI.

The early work of using H_∞ optimization techniques for robust FDI was based on the use of the factorization approach (Ding and Frank 1990; Ding *et al.* 2000). The factorization-based H_∞ optimization technique is useful in solving FDI problems. However, the more elegant and advanced H_∞ optimization methods are based on the use of the Algebraic Riccati Equation (ARE) (Zhou *et al.* 1996). Mangoubi *et al.* (1992) first solved the robust FDI estimation problem using the ARE approach via the use of H_∞ and μ robust estimator synthesis methods developed by Appleby *et al.* (1991). A direct formulation of the FDI problem as a robust H_∞ filter design problem with the solution of an ARE was given in Edelmayer *et al.* (1997). To deal with modeling errors, as well as disturbances in robust FDI design, in Niemann and Stoustrup (1996), the concept of modeling error blocks into the standard H_∞ observer design was introduced. The weighting factors are then introduced in the problem formulation for finding an optimal FDI solution. This is further extended to nonlinear systems where the nonlinearity is treated in the same way as a modeling error block (Stoustrup and Niemann 1998; Stoustrup *et al.* 1997).

The majority of studies discussed so far involve the use of a slightly modified H_∞ filter for the residual generation, that is, the design objective is to minimize the effect of disturbances and modeling errors on the estimation error and subsequently, on the residual. However, robust residual generation is different from the robust estimation because it does not only require the disturbance attenuation. The residual has to remain sensitive to faults while the effect of disturbance is minimized. In Sauter *et al.* (1997), the problem where the fault sensitivity is enhanced by applying an optimal post-filter to the "primary residual" was studied. The problem of enhancing fault sensitivity while increasing robustness against disturbances and modeling errors was studied extensively in Sadrnia *et al.* (1997). The essential idea is to reach an acceptable compromise between disturbance robustness and fault sensitivity. In the beginning, an observer with a very small disturbance sensitivity bound is designed via an ARE. Then, the fault sensitivity is checked. If the fault sensitivity is too small, the disturbance robustness requirement should be relaxed, that is, to design another optimal observer with an increased disturbance sensitivity bound. This procedure is likely to be repeated several times. The final goal is to find a design that provides the maximum ratio between fault sensitivity and disturbance sensitivity.

In Chen and Patton (1999, 2000), the robust residual generation problem was formulated within the standard H_∞ filtering framework, that is, to generate the residual whose sensitivity to disturbances is minimized. To facilitate reliable FDI, the residual sensitivity to faults has to be maintained (or maximized) in addition to the minimization of the disturbance sensitivity. This problem was solved via the minimization of the difference between the residual and the fault against the disturbance and the fault, that is, the objective is to replicate the fault using the residual. In this way, the residual sensitivity to the fault is indirectly maximized. The residual sensitivity to the modeling error can be minimized if the modeling error is approximately represented by the disturbance vector with the estimated distribution matrix (Chen and Patton 1999). However, the modeling error can be handled directly using standard H_∞. In Chen and Patton (1999), the way of including the modeling error in the robust residual design within the standard H_∞ framework was explained.

Generally speaking, the robust FDI approach can be approached in different ways. It is therefore important to mention the design principle of residual generators under a certain performance index (Basseville 1997; Frank *et al.* 2000). This is indeed a reasonable extension of the unknown input residual generator design, in which, instead of full decoupling, a compromise between the robustness and sensitivity is made. It is worth focusing the attention to this scheme, due to its important role in theoretical studies and its relationship to the residual evaluation and integrated design of FDI systems. Since the goal of residual generation is to enhance the robustness of the residual to the model uncertainty without loss of its sensitivity to the faults, the minimization of performance index in the form of equation [1.68] (Frank *et al.* 2000):

$$J = \frac{\|\frac{\partial r}{\partial d}\|}{\|\frac{\partial r}{\partial f}\|} \text{ or } J = \|\frac{\partial r}{\partial d}\| \text{ with } \|\frac{\partial r}{\partial f}\| > \alpha \qquad [1.68]$$

is widely recognized as a suitable design objective. Associated with the norm used, the type of the residual generator and the mathematical tool adopted, a number of optimization approaches have been developed (Frank *et al.* 2000). In Ding *et al.* (2000), a unified solution for a number of optimization problems was derived, and thus, a satisfactory solution to the above-defined optimization problem was provided 10 years after it was first proposed. In Frank *et al.* (2000), a brief review of the state of art of the solutions can be found, while (Hou and Patton 1996, 1997; Frank *et al.* 2000) first introduced the H_∞/H_- method. According to the norm selected, by minimizing the performance index of equation [1.68] over a specified range, an approximate decoupling design can be achieved (Ding and Frank 1990; Patton and Hou 1997; Frank and Ding 1997; Ding *et al.* 1999). Moreover, the approximated design for optimal disturbance decoupling can also be solved in the time domain (Wünnenberg 1990; Chen *et al.* 1993). On the other hand, with reference to the modeling errors, the robust problem is more difficult to solve.

Two main techniques have been described in Patton and Chen (1994b) and Chen *et al.* (1996). In the first case, the uncertainty is taken into account at the residual design stage (Chen *et al.* 1996); this is known as *active robustness* in fault diagnosis (Patton and Chen 1994b). The active way of achieving a robust solution is to approximate uncertainties, that is, represent approximate modeling errors as disturbances (Chen and Patton 1999). In this case, it is assumed an unknown vector, while its effect is estimated transfer function. When this approximate structure is exploited to design disturbance decoupling residual generators, robust FDI can be achieved. An active approach applied to the robust active fault tolerance will be addressed in the following chapters.

The second approach called *passive robustness* makes use of a residual evaluator with adaptive threshold. As a simple example, it is assumed that the residual generation uncertainty is only represented by modeling errors. Once the fault-free residual r has been designed, under the assumption that the modeling errors are bounded by a value δ, an adaptive threshold $\varepsilon(t)$ can be generated by an adaptive system.

In this case, the threshold $\varepsilon(t)$ is no longer fixed, but depends on the input $u(t)$, thus being adaptive to the system operating point. A robust FDI technique with the threshold adaptor or selector was originally presented in Clark (1989) and Emami-Naeini *et al.* (1988). This method represents a passive approach, since no effort is made to design a robust residual. Even if disturbance decoupling methods for robust FDI have been studied extensively, their effectiveness regarding real problems has not been fully demonstrated. The main difficulty arises as most of the disturbances only account for a small percentage of the uncertainty in the real system. The presented disturbance decoupling methods cannot be applied directly to the systems with other uncertainties, such as modeling errors. The estimation and approximate representation of modeling errors, as well as other uncertain factors such as the disturbance term, provides a practical way to tackle the robustness issue for technical processes, as shown in this book.

Some concluding remarks can finally be drawn here. This chapter provided some theoretical study results for the detection and diagnosis of faults in the actuators, systems and sensors of dynamic processes and technical systems, through the use of different FDD schemes. Residual generators were designed from the nonlinear input–output descriptions of the system under diagnosis, and the disturbance decoupling was obtained. Procedures for optimizing the residual generator fault sensitivity and dynamic response were also suggested. An important aspect of the strategies based on residual generators is the simplicity of the technique used to generate these residuals when compared with different schemes. The algorithmic simplicity is a very important aspect when considering the need for verification and validation of demonstrable schemes that will be applied to real processes. The more complex the computations required to implement the scheme, the higher the cost and complexity in terms of verification and validation.

On the other hand, nonlinear methodologies rely on a design scheme, based on the structural decoupling of the disturbance obtained by means of a coordinate transformation in the state space and in the output space. To apply the nonlinear theory, simplified models of the system under investigation may be required. The mixed $\mathcal{H}_-/\mathcal{H}_\infty$ optimization of the trade-off between fault sensitivity, disturbances and modeling errors is now well understood in the theoretical work and is a promising area for application study to dynamic processes and technical systems.

The nonlinear fault diagnosis strategies were also based on the adaptive filters scheme. In addition to a proper detection and isolation, these methods also provided a fault size estimation. This feature is unusual for fault detection and isolation methods and can be fundamental for the online automatic control system reconfiguration, in order to recover a faulty operating condition. Compared with similar methods proposed in the literature, the nonlinear adaptive fault diagnosis technique described here has the advantage of being applicable to more general classes of nonlinear systems and less sensitive to measurement noise, since it does not use input–output signal derivatives. Suitable filtering algorithms for stochastic systems were also recalled. The knowledge regarding the noise process acting on the system under diagnosis can be exploited by the fault diagnosis method design, hence, the proposed scheme provides a possible solution to nonlinear system diagnosis with non-Gaussian noise and disturbance. The main advantage of nonlinear-based FDD techniques with disturbance decoupling features is represented by the fact that they directly take into account the model nonlinearity and the system reality-model mismatch.

The fault diagnosis techniques that have been outlined in this chapter can be applied to high-fidelity simulators of real dynamic processes and technical systems, which are able to take into account disturbances and measurement errors acting on the systems under investigation. Moreover, the robustness characteristics and the achievable performances of the fault diagnosis approaches described have been carefully considered and investigated. The effectiveness of the proposed diagnosis schemes was shown by simulations and a comparison with widely used data-driven and model-based FDI and an FDD scheme with disturbance decoupling. The reliability and the robustness properties of the designed residual generators to model uncertainty, disturbances and measurements noise will be analyzed via extensive simulations, including the use of Monte Carlo simulation experiments to tune the FDI and FDD parameters.

Finally, the need to bridge the design gap between FDD and recovery mechanisms, that is, the sustainable schemes, can be obtained when fault diagnosis and fault-tolerant control strategies can be properly combined.

1.7. Summary

This chapter presented a tutorial on the basic principles of model-based and data-driven FDI, with application to dynamic systems and technical processes.

The FDI problem has been formalized in a uniform framework by presenting a mathematical description and definition. Within this framework, the residual generation has been identified as a central issue in both model-based and data-driven FDI. By choosing the proper design approach, the FDI task can be performed.

The residual generator was summarized in different residual generation structures. The ways of designing residuals for fault isolation were also discussed. The most commonly used residual generation strategies were recalled by presenting related problems and discussing the applicability of model-based and data-driven FDI methods to dynamic processes, such as dynamic processes, and technical systems.

It is worth noting that the success of fault diagnosis depends on the quality of the residuals. Successful diagnosis requires residual signals that should be robust with respect to modeling uncertainty. The robust FDI problem was also discussed in this chapter and the implementation of robust residual generators will be shown in the following chapters of the book.

Finally, data-driven FDI methods such as fuzzy logic, qualitative modeling and NNs were also considered and the concept of integrated knowledge-based fault diagnosis, utilizing both analytical and heuristic information, was illustrated.

1.8. References

Appleby, B., Dowdle, J., Vander Velde, W. (1991). Robust estimator design using μ synthesis. In *Proc. of the 30th Conf. on Decision & Control*, Brighton, UK, 640–644.

Ayoubi, M. (1995). Neuro–fuzzy structure for rule generation and application in the fault diagnosis of technical processes. In *Proc. of the American Control Conference, ACC'95*, Washington. IEEE, Piscataway, NJ, 2757–2761.

Babuka, R. (1998). *Fuzzy Modeling for Control*. Kluwer Academic Publishers, Boston, MA.

Babuka, R. (2012). *Fuzzy Modeling for Control*, vol. 12. Springer Science & Business Media, Berlin, Heidelberg.

Basseville, M. (1988). Detecting changes in signals and systems: A survey. *Automatica*, 24(3), 309–326.

Basseville, M. (1997). Information criteria for residual generation and fault detection and isolation. *Automatica*, (33), 783–803.

Basseville, M. and Benveniste, A. (1986). Detection of abrupt changes in signals and dynamical systems. In *Lecture Notes in Control and Information Sciences*, vol. 77. Springer-Verlag, London.

Basseville, M. and Nikiforov, I.V. (1993). *Detection of Abrupt Changes: Theory and Application*. Prentice–Hall Inc, Upper Saddle River, NJ.

Beard, R.V. (1971). Failure accomodation in linear systems through self-reorganisation. Technical Report MVT-71-1, Man Vehicle Lab., Cambridge, MA.

Beghelli, S., Guidorzi, R.P., Soverini, U. (1990). The Frisch scheme in dynamic system identification. *Automatica*, 26(1), 171–176.

Bezdek, J.C. (2013). *Pattern Recognition with Fuzzy Objective Function Algorithms*. Springer Science & Business Media, Berlin, Heidelberg.

Bonfè, M., Castaldi, P., Geri, W., Simani, S. (2006). Fault detection and isolation for on-board sensors of a general aviation aircraft. *International Journal of Adaptive Control and Signal Processing*, 20(8), 381–408.

Boyd, S., Ghaoui, L., Feron, E., Balakrishnan, V. (1994). *Linear Matrix Inequalities in System and Control Theory*. SIAM, Philadelphia.

Brown, M. and Harris, C. (1994). *Neurofuzzy Adaptive Modelling and Control*. Prentice Hall, Upper Saddle River. NJ.

Calado, J., Korbicz, J., Patan, K., Patton, R., Sá da Costa, J. (2001). Soft computing approaches to fault diagnosis for dynamic systems. *European Journal of Control*, 7(2–3), 248–286.

Carpenter, G. and Grossberg, S. (1987). A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics and Image Processing*, 37, 54–115.

Chen, J. and Patton, R.J. (1999). *Robust Model-Based Fault Diagnosis for Dynamic Systems*. Kluwer Academic Publishers, Boston, MA.

Chen, J. and Patton, R.J. (2000). Standard H_∞ filter formulation of robust fault detection. In *SAFEPROCESS2000, 4th IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes, vol. 1, IFAC 2000*, Budapest, Hungary, Edelmayer, A.M. (ed.). Elsevier Ltd., London.

Chen, J. and Patton, R.J. (2001). Fault-tolerant control systems design using the linear matrix inequality method. In *European Control Conference, ECC'01*, Porto, Portugal. IEEE, Piscataway, NJ, 1993–1998.

Chen, S., Billings, A.S., Cowan, C.F.N., Grant, P.M. (1990). Practical identification of NARMAX models using radial basis function. *Int. J. Control*, 52, 1327–1350.

Chen, J., Patton, R.J., Zhang, H.Y. (1993). A multi-criteria optimization approach to the design of robust fault detection algorithm. In *Proc. of Int. Conf. on Fault Diagnosis: TOOLDIAG'93*, Toulouse, France. Elsevier Ltd., London, 783–803.

Chen, J., Patton, R.J., Zhang, H.Y. (1996). Design of unknown input observer and robust fault detection filters. *Int. J. Control*, 63(1), 85–105.

Chiang, L.H., Russel, E.L., Braatz, R.D. (2001). *Fault Detection Diagnosis in Industrial Systems*. Springer-Verlag London Limited, London.

Chow, E.Y. and Willsky, A.S. (1980). Issue in the development of a general algorithm for reliable failure detection. In *Proc. of the 19th Conf. on Decision & Control*, Albuquerque, NM. IEEE, Piscataway, NJ.

Chow, E.Y. and Willsky, A.S. (1984). Analytical redundancy and the design of robust detection systems. *IEEE Trans. Automatic Control*, 29(7), 603–614.

Chung, W.H. and Speyer, J.L. (1998). A game theoretic fault detection filter. *IEEE Trans. on Automatic Control*, 43(2), 143–161.

Clark, R.N. (1978). Instrument fault detection. *IEEE Trans. Aero. & Electronic Systems*, 14(3), 456–465.

Clark, R.N. (1989). Model-based methods for fault diagnosis: Some guide-lines. In *Fault Diagnosis in Dynamic Systems: Theory and Application*, Patton, R., Clark, R., Frank, P.M. (eds). Prentice Hall, Upper Saddle River, NJ.

Daly, K.C., Gai, E., Harrison, J.V. (1979). Generalized likelihood test for FDI in redundancy sensor configurations. *J. of Guidance, Control & Dynamics*, 2(1), 9–17.

Davis, L. (1991). *Handbook of Genetic Algorithms*. Van Nostrand Reinhold, New York.

De Persis, C. and Isidori, A. (2000). On the observability codistributions of a nonlinear system. *Systems and Control Letters*, 40, 297–304.

De Persis, C. and Isidori, A. (2001). A geometric approach to nonlinear fault detection and isolation. *IEEE Transactions on Automatic Control*, 46(6), 853–865.

Delmaire, G., Cassar, P., Staroswiecki, M., Christophe, C. (1999). Comparison of multivariable identification and parity space techniques for FDI purpose in M.I.M.O. systems. *Proceedings of the 1999 European Control Conference (ECC)*, Karlsruhe, Germany. IEEE, Piscataway, NJ, 2789–2794.

Dexter, A.L. and Benouarets, M. (1997). Model-based fault diagnosis using fuzzy matching. *IEEE Trans. on Sys. Man. and Cyber. Part A: Sys. & Humans*, 27(5), 673–682.

Dietz, W.E., Kiech, E.L., Ali, M. (1989). Jet and rocket engine fault diagnosis in real time. *J. of Neural Network Computing*, 1, 5–18.

Ding, X. and Frank, P.M. (1990). Fault detection via factorization approach. *Syst. Contr. Lett.*, 14(5), 431–436.

Ding, S.X., Jeinsch, T., Ding, E.L., Zhou, D., Wang, G. (1999). Application of observer-based FDI schemes to the three tank system. *Proceedings of the 1999 European Control Conference (ECC)*, Karlsruhe, Germany. IEEE, Piscataway, NJ, 4438–4443

Ding, S.X., Jeinsch, T., Frank, P.M., Dind, E.L. (2000). A unified approach to the optimisation of fault detection systems. *Int. J. of Adaptive Control and Signal Processing*, 14(7), 725–745.

Duan, G., How, D., Patton, R. (2002). Robust fault detection in descriptor systems via generalised unknown input observers. *Int. J. Systems Science*, 32(2), 7724–7729.

Edelmayer, A., Bokor, J., Keviczky, L. (1997). A scaled L_2 optimisation approach for improving sensitivity of $\mathrm{H}_{i}nfty$ detection filters for LTV systems. In *Preprints of the 2nd IFAC Symp. on Robust Control Design: RECOND97*, Budapest, Hungary, Bányász, C. (ed.). Elsevier, London.

Emami-Naeini, A., Akhter, M., Rock, M. (1988). Effect of model uncertainty on failure detection: The threshold selector. *IEEE Trans. on Automatic Control*, 33(12), 1105–1115.

Fantuzzi, C. and Rovatti, R. (1996). On the approximation capabilities of the homogeneous Takagi-Sugeno model. In *Proceedings of the 5th IEEE International Conference on Fuzzy Systems*. IEEE, Piscataway, NJ, 2, 1067–1072.

Fantuzzi, C., Simani, S., Beghelli, S., Rovatti, R. (2002). Identification of piecewise affine models in noisy environment. *International Journal of Control*, 75(18), 1472–1485.

Frank, P.M. (1993). Advances in observer-based fault diagnosis. *Proc. TOOLDIAG'93 Conference*, CERT, Toulose, France, 3, 817–836.

Frank, P.M. (1994). Enhancement of robustness on observer-based fault detection. *International Journal of Control*, 59(4), 955–983.

Frank, P.M. and Ding, X. (1997). Survey of robust residual generation and evaluation methods in observer-based fault detection system. *Journal of Process Control*, 7(6), 403–424.

Frank, P.M., Ding, S.X., Köpper-Seliger, B. (2000). Current developments in the theory of FDI. In *SAFEPROCESS'00: Preprints of the IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes*, vol. 1, Budapest, Hungary. Elsevier Ltd., London, 16–27.

Füssel, D., Ballé, P., Isermann, R. (1997). Closed loop fault diagnosis based on a non-linear process model and automatic fuzzy rule generation. In *Proc. of IFAC Symposium on Fault Detection, Supervision and Safety for Technical Process SAFEPROCESS'97*, The University of Hull. Elsevier Ltd., London.

Gertler, J. (1991). Generating directional residuals with dynamic parity equations. *Proc. IFAC/IMACS Symp. SAFEPROCESS'91*. 26(2), 507–512.

Gertler, J. (1998). *Fault Detection and Diagnosis in Engineering Systems*. Marcel Dekker, New York.

Gertler, J. and Monajemy, R. (1993). Generating directional residuals with dynamic parity equations. *Proc. of the 12th IFAC World Congress*, 26(2), 505–510.

Gertler, J. and Singer, D. (1990). A new structural framework for parity equation-based failure detection and isolation. *Automatica*, 26(2), 381–388.

Graaff, A.J. and Engelbrecht, A.P. (2012). Clustering data in stationary environments with a local network neighborhood artificial immune system. *International Journal of Machine Learning and Cybernetics*, 3(1), 1–26.

Hadamard, J. (1964). *La théorie des équations aux derivées partielles*. Pékin, Éditions Scientifiques, Paris.

Hagan, M.T. and Menhaj, M.B. (1994). Training feedforward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks*, 5(6), 989–993.

Haykin, S.S. (2009). *Neural Networks and Learning Machines*, vol. 3. Pearson Education, Upper Saddle River, NJ.

Himmelblau, D.M. (1978). *Fault Diagnosis in Chemical and Petrochemical Processes*. Elsevier, Amsterdam.

Himmelblau, D.M., Barker, R.W., Suewatanakul, W. (1991). Fault classification with the aid of artificial neural networks. In *IFAC/IMACS Symposium SAFEPROCESS '91*, Baden Baden, Germany. Elsevier Ltd., London, 2, 369–373.

Hoskins, J.C. and Himmelblau, D.M. (1988). Artificial neural network models of knowledge representation in chemical engineering. *Comp. Chem. Eng.*, 12, 881–890.

Hou, M. and Patton, R.J. (1996). An LMI approach to H_{-}/H_{∞} fault detection observers. In *Proceedings of the 1996 UKACC International Conference on Control'96*. IEEE, Piscataway, NJ, 305–310.

Hou, M. and Patton, R.J. (1997). An H_{∞}/H_{-} approach to the design of robust fault diagnosis observers based upon LMI optimisation. *Proceedings of the 4th European Control Conference. ECC'97*, Brussels. Wiley, Hoboken, NJ.

Isermann, R. (1984). Process fault detection based on modeling and estimation methods: A survey. *Automatica*, 20(4), 387–404.

Isermann, R. (1992). Estimation of physical parameters for dynamic processes with application to an industrial robot. *Int. J. of Control*, 55, 1287–1298.

Isermann, R. (1994). Integration of fault detection and diagnosis methods. *Proc. IFAC SAFEPROCESS Symposium '94*, Espoo, Finland. Elsevier Ltd., London, 27(5), 575–590.

Isermann, R. (1997). Supervision, fault detection and fault diagnosis methods: An introduction. *Control Engineering Practice*, 5(5), 639–652.

Isermann, R. (1998). On fuzzy logic applications for automatic control, supervision and fault diagnosis. *IEEE Trans. on Sys. Man. and Cyber. Part A: Sys. & Humans*, 28(2), 221–235.

Isermann, R. (2005). *Fault–Diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance*, 1st edition. Springer-Verlag, Weinheim, Germany.

Isermann, R. and Ballé, P. (1997). Trends in the application of model-based fault detection and diagnosis of technical processes. *Control Engineering Practice*, 5(5), 709–719.

Jain, A.K., Murty, M.N., Flynn, P.J. (1999). Data clustering: A review. *ACM Computing Surveys (CSUR)*, 31(3), 264–323.

Jang, J.S.R. (1993). ANFIS: Adaptive-Network-based Fuzzy Inference System. *IEEE Transactions on Systems, Man., & Cybernetics*, 23(3), 665–684.

Jang, J.S.R. (1994). Structure determination in fuzzy modelling: A fuzzy CART approach. *Proc. of IEEE International Conf. on Fuzzy Systems*, Orlando, FL. IEEE, Piscataway, NJ, 1, 480–485.

Jang, J.S.R. and Sun, C.T. (1995). Neuro-fuzzy modeling and control. *Proc. IEEE*, 83(3), 378–405.

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

Jun, W., Shitong, W., Chung, F.-l. (2011). Positive and negative fuzzy rule system, extreme learning machine and image classification. *International Journal of Machine Learning and Cybernetics*, 2(4), 261–271.

Kalman, R.E. (1982a). Identification from real data. In *Current Developments in the Interface: Economics, Econometrics, Mathematics*, Hazewinkel, M., Rinnoy Kan, A.H.G. (eds). D. Reidel, Dordrecht, The Netherlands.

Kalman, R.E. (1982b). System identification from noisy data. In *Dynamical System II*, Bednarek, A.R., Cesari, L. (eds). Academic Press, New York.

Kalman, R.E. (1984). Identification of noisy systems. *50th Anniversary Symp. Steklov Institute of Mathematics*. U.S.S.R. Academy of Sciences, Moscow.

Kalman, R.E. (1990). Nine lectures on identification. In *Lecture Notes in Economics and Mathematical Systems*, Springer-Verlag, Berlin.

Korbicz, J., Patan, K., Obuchowicz, A. (1999). Dynamic neural network for process modelling in fault detection and isolation systems. *Applied Mathematics and Computer Science*, 9(2), 519–546.

Leontaritis, I. and Billings, S.A. (1985a). Input-output parametric models for non-linear systems part II: Stochastic non-linear systems. *Int. J. Control*, 41(2), 329–344.

Leontaritis, I. and Billings, S.A. (1985b). Input-output parametric models for non-linear systems part I: Deterministic non-linear systems. *Int. J. Control*, 41(2), 303–328.

Liu, G.P. (2012). *Nonlinear Identification and Control: A Neural Network Approach*. Springer Science & Business Media, Berlin.

Liu, G.P. and Patton, R.J. (1998). *Eigenstructure Assignment for Control System Design*. John Wiley & Sons, England.

Ljung, L. (1999). *System Identification: Theory for the User*, 2nd edition. Prentice Hall, Englewood Cliffs, NJ.

Lou, X., Willsky, A., Verghese, G. (1986). Optimal robust redudancy relations for failure detection in uncertainty systems. *Automatica*, 22(3), 333–344.

Mamdani, E. (1976). Advances in the linguistic synthesis of fuzzy controllers. *Int. J. Man-Machine Studies*, 8, 669–678.

Mamdani, E. and Assilian, S. (1995). An experiment in linguistic synthesis with fuzzy logic controller. *Int. J. Man-Machine Studies*, 7(1), 1–13.

Mangoubi, R., Appleby, B.D., Farrell, J.R. (1992). Robust estimation in fault detection. In *Proc. of the 31st Conf. on Decision & Control*, Tucson, AZ. IEEE, Piscataway, NJ, 2317–2322.

Marquardt, D.W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2), 431–441.

Massoumnia, M.A. (1986). A geometric appoach to failure detection and identification in linear systems. PhD Thesis, Massachusetts Institute of Technology, Cambridge, MA.

Massoumnia, M., Verghese, G.C., Willsky, A.S. (1989). Failure detection and identification. *IEEE Trans. Automat. Contr.*, 34, 316–321.

McDuff, R.J. and Simpson, P.K. (1990). An adaptive reasonance diagnostic system. *J. of Neural Network Computing*, 2, 19–29.

Medsker, L. and Jain, L.C. (1999). *Recurrent Neural Networks: Design and Applications*. CRC Press, Cleveland, OH.

Meneganti, M., Saviello, F., Tagliaferri, R. (1998). Fuzzy neural networks for classification and detection of anomalies. *IEEE Trans. on Neural Networks*, 9(5), 848–861.

Morozov, V. (1984). *Methods for Solving Incorrectly Posed Problems*. Springer, Berlin.

Nelles, O. (2001). *Nonlinear System Identification*. Springer-Verlag, Berlin, Heidelberg.

Nelles, O. and Isermann, R. (1996). Basis function networks for interpolation of local linear models. In *Proc. of the 35th IEEE Conference on Decision and Control*, Kobe, Japan. IEEE, Piscataway, 4, 470–475.

Niemann, H. and Stoustrup, J. (1996). Filter design for failure detection and isolation in the presence of modelling errors and disturbances. In *Proc. of the 35th IEEE Conf. on Decision and Contr.*, Kobe, Japan. IEEE, Piscataway, 1155–1160.

Norgaard, M., Poulsen, N.K., Ravn, O. (2000). New developments in state estimation for nonlinear systems. *Automatica*, 36, 1627–1638.

Patton, R.J. and Chen, J. (1991). A review of parity space approaches to fault diagnosis. *IFAC Symposium SAFEPROCESS '91*, Baden-Baden, Germany. Elsevier Ltd., London, 24(6), 65–81.

Patton, R.J. and Chen, J. (1993). Optimal selection of unknown input distribution matrix in the design of robust observers for fault diagnosis. *Automatica*, 29(4), 837–841.

Patton, R.J. and Chen, J. (1994a). A review of parity space approaches to fault diagnosis for aerospace systems. *AIAA Journal of Guidance, Control & Dynamics*, 17(2), 278–285.

Patton, R.J. and Chen, J. (1994b). A review of parity space approaches to fault diagnosis for aerospace systems. *AIAA J. of Guidance, Contr. & Dynamics*, 17(2), 278–285.

Patton, R.J. and Chen, J. (1997). Observer-based fault detection and isolation: Robustness and applications. *Control Eng. Practice*, 5(5), 671–682.

Patton, R.J. and Chen, J. (2000). On eigenstructure assignment for robust fault diagnosis. *Int. J. of Robust & Non-Linear Control*, 10(9), 1193–1208.

Patton, R.J. and Hou, M. (1997). H_∞ estimation and robust fault detection. *Proc. of the 1997 European Control Conference, ECC'97*, Brussels, Belgium. Elsevier Ltd., London, 2520–2524.

Patton, R.J., Frank, P.M., Clark, R.N. (eds) (1989). *Fault Diagnosis in Dynamic Systems, Theory and Application*. Prentice Hall, London.

Patton, R.J., Lopez-Toribio, C.J., Uppal, F.I. (1999). Artificial intelligence approaches to fault diagnosis. *Applied Mathematics and Computer Science*, 9(3), 471–518.

Patton, R.J., Frank, P.M., Clark, R.N. (eds) (2000). *Issues of Fault Diagnosis for Dynamic Systems*. Springer-Verlag, London.

Pau, L.F. (1981). *Failure Diagnosis and Performance Control*. Marcel Dekker, New York.

Potter, I.E. and Suman, M.C. (1977). Thresholdless redundancy management with array of skewed instruments. Technical Report AGARDOGRAPH-224, AGARD, Integrity in Electronic Flight Control Systems [Online]. Available at: https://jglobal.jst.go.jp/en/detail?JGLOBAL_ID=201002064777220137.

Ray, A. and Luck, R. (1991). An introduction to sensor signal validation in redundant measurement systems. *IEEE, Contr. Syst. Mag.*, 11(2), 44–49.

Rovatti, R. (1996). Takagi-Sugeno models as approximators in Sobolev norms: The SISO case. In *Proceedings of the Fifth IEEE International Conference on Fuzzy Systems*. IEEE, Piscataway, NJ, 2, 1060–1066.

Sadrnia, M.A., Chen, J., Patton, R.J. (1997). Robust H_∞/μ observer-based residual generation for fault diagnosis. In *Proc. of the IFAC Symp. on Fault Detection, Supervision and Safety for Technical Processes: SAFEPROCESS'97*, University of Hull. Pergamon, Oxford, 155–162.

Sauter, D., Rambeaux, F., Hamelin, F. (1997). Robust fault diagnosis in a H_∞ setting. In *Proc. of the IFAC Symp. on Fault Detection, Supervision and Safety for Technical Processes: SAFEPROCESS'97*, University of Hull. Pergamon, Oxford, UK, 867–874.

Simani, S., Fantuzzi, C., Rovatti, R., Beghelli, S. (1999). Parameter identification for piecewise-affine fuzzy models in noisy environment. *International Journal of Approximate Reasoning*, 22(1), 149–167.

Simani, S., Fantuzzi, C., Patton, R.J. (2003). *Model-based Fault Diagnosis in Dynamic Systems Using Identification Techniques*, 1st edition. Springer-Verlag, London.

Söderström, T. and Stoica, P. (1987). *System Identification*. Prentice Hall, Englewood Cliffs, NJ.

Speyer, J.L. (1999). Residual sensitive fault detection filters. In *7th IEEE Mediterranean Conference on Control and Automation (MED'99)*, Haifa, Israel. IEEE, Piscataway, NJ, 1962–1973.

Stoustrup, I. and Niemann, H. (1998). Fault detection for nonlinear systems–A standard problem approach. In *Proc. of the 37th IEEE Conf. on Decision & Control*, Tampa, FL. IEEE, Piscataway, NJ, 96–101.

Stoustrup, J., Grimble, M.J., Niemann, H. (1997). Design of integrated systems for the control and detection of actuator/sensor faults. *Sensor Review*, 17(2), 138–149.

Sugeno, M. and Kang, G. (1988). Structure identification of fuzzy model. *Fuzzy Set and Systems*, 28, 15–33.

Tachibana, K. and Furuhashi, T. (1994). A hierarchical fuzzy modelling method using genetic algorithm for identification of concise submodels. *Proc. of 2nd Int. Conference on Knowledge-Based Intelligent Electronic Systems*, Adelaide, Australia. IEEE, Piscataway, NJ.

Takagi, T. and Sugeno, M. (1985a). Fuzzy identification of systems and its application to modeling and control. *IEEE Transaction on System, Man and Cybernetics*, 15(1), 116–132.

Takagi, T. and Sugeno, M. (1985b). Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man and Cybernetics*, (1), 116–132.

Tikhonov, A. and Arsenin, V. (1977). *Solution of Ill-posed Problems*. Winston and Wiley, Washington.

Tou, J.T. and Gonzalez, R.C. (1974). *Pattern Recognition Principles*. Addison-Wesley Publishing, Boston, MA.

Venkatasubramanian, V. and Chan, K. (1989). A neural network methodology for process fault diagnosis. *AIChE J.*, 35, 1993–2002.

Willsky, A.S. (1976). A survey of design methods for failure detection in dynamic systems. *Automatica*, 12(6), 601–611.

Wu, Z.Q. and Harris, C.J. (1996). Neuro-fuzzy modelling and state estimation. In *IEEE Medit. Symp. on Control and Automation: Circuits, Systems and Computers '96*. Hellenic Naval Academy, Piraeus, Greece.

Wünnenberg, J. (1990). Observer-based fault detection in dynamic systems. PhD Thesis, University of Duisburg, Duisburg, Germany.

Wünnenberg, J. and Frank, P.M. (1987). Sensor fault detection via robust observer. In *System Fault Diagnosis, Reliability, and Related Knowledge-Based Approaches*, Tzafestas, S.G., Singh, M., Schmidt, G. (eds). Springer, London, UK.

Zhou, K., Doyle, J.C., Glover, K. (1996). *Robust and Optimal Control*. Prentice Hall, Englewood Cliffs, NJ.

Zolghadri, A. (1996). An algorithm for real-time failure detection in Kalman filters. *IEEE Transactions on Automatic Control*, 41(10), 1537–1539.